AUTONOMA UNIVERSITY OF MADRID

MOLECULAR BIOLOGY DEPARTMENT

Identification and characterization of cancer initiating cells in K-Ras driven Non Small Cell Lung Cancer

DOCTORAL THESIS

Mainardi Sara

Madrid, 2013

Molecular Biology Department

Faculty of Science

AUTONOMA UNIVERSITY OF MADRID

Identification and characterization of cancer initiating cells in

K-Ras driven Non Small Cell Lung Cancer

Sara Mainardi

Graduated in Biotechnology

Director: Dr. Mariano Barbacid Montalbán

Spanish National Cancer Center (CNIO)

Madrid, 2013

ACKNOWLEDGEMENTS

The completion of this doctoral thesis is due to the opportunity given to me by my supervisor Dr. Mariano Barbacid, who I would like to heartily thank for all his support and belief in me. I would also like to thank my academic tutor Dr. Eva Richard. Although she only accompanied me during this last year it was a pleasure to count on such a nice and available person. And also I would like to thank all the members of my PhD Committe, Dr. Eduard Batllé, Dr. Christopher Heeschen and Dr. Erwin Wagner for the useful scientific discussion during our meetings and for always being supportive.

One of the first scientific life lessons my boss taught me is that “the most expensive thing in science is time”. Well, I think in times like this, this is not completely true, so I would like to thank La Caixa foundation as well for financially support the work presented in this thesis.

Finalmente, quiero agradecer a todas las personas, compañeros y amigos, que han hecho posible que en estos años aprendiera mucho más que como desarrollar un proyecto de investigación cientifica. Este tiempo en el CNIO, en Madrid, en España, ha sido una inolvidable y muy enriquecedora experiencia de vida. Lo malo de este trabajo, que tanta movilidad conlleva, es que acabas viendo muchos de tus amigos y compañeros irse del labo más pronto de lo que te gustaría. Lo bueno, es que acabas con amigos en todo el mundo, y esto hace que el mundo sea todo como tu casa. Cuando empecé en Oncología Experimental me acogieron Sarah, Rafa, Javier, Marta, Antonio, Mikel, Alberto, Carolina, Matthias, Fonchito…y qué pocos quedan ahora de este antiguo y tan agradable grupo, al que le agradezco de corazón haber facilitado la integración de esta pobre espaguetti recién llegada del otro lado del Mediterranéo. De algunos me ha dado tiempo hasta a verlos volver, verdad Chumi? Que alegría tener de nuevo a alguien que valore la combinación de los colores en la ropa, y en particular el naranja. Un gracias especial a Sarah, fuiste la primera que fue amable conmigo y nunca dejaste de serlo, realmente mi primer castellano lo aprendí de una francesa! Conocer a una persona tan buena y sincera merece 5 años de tesis y mucho más!

Y si es verdad que el grupo ha cambiado mucho desde que llegué aquí, también es cierto que los substitutos han sido más que dignos, y afortunadamente conseguimos entre todos que se pasen ratos buenos en el labo, a pesar de los westerns que de repente dejan de salir, de la reverse transcriptase que siempre se acaba, y de los recuentos de cubetas que nos producen pesadillas. Iré por orden de pasillo. Lucía, Malaguita, gracias por llevar cada día un poco de color al labo, para mi no eres perroflauta, eres guapa por dentro y por fuera. Y gracias por no llevarme la serie completa de Twin Peaks, sino estaría todavía con Materiales y Metodos. Magdi, no hace falta que escriba en inglés pues ya te escuché hablar español y si que aprendiste! Gracias por tu amabilidad, te deseo todo lo mejor con tu tesis, ánimo! Harrys, tu si que todavía estás autorizado a no entender el castellano, peró seguro que vas a entender que estamos todos muy contentos de tenerte por aquí, me encanta tu curiosidad y tus cuentos de la India, espero un día poder visitar tu pays pues me fascina mucho! Isabel, aunque no hemos

tenido mucho trato te agradezco tu disponibilidad cada vez que lo he necesidado…eres la gran experta del FACS y siempre que tendré que combinar 10 o más marcadores te preguntaré a ti! Sarita…ya te he agradecido, no puedo siempre agradecerte a ti! (es broma, sabes que te quiero mucho!). Matthias, gracias por representar la seriedad en este labo de chillones! Un día me contarás como pudiste mandar un paper en tan solo una tarde, algún truco tiene que haber, y yo desde luego no lo he pillado. Carolina, a ti te tengo que dar la gracias por muchas cosas (y no me refiero al buffer fosfato y a la X-Gal…bueno a esto también). Gracias por tu alegría y también por tus enfados pues las dos cosas le dan vida al labo! Gracias por ser una buena amiga, por preocuparte por mi, por este viaje tan chulo a San Francisco que si cierro los ojos no veo al Golden Bridge sino a un autobús de gente bailando y tomando alcohol a las 9 de la mañana…y nosotras con nuestros cafés…Cathy, nuestra Americana, tu eres diferente, tu por la mañana te tomas café (o esta cosa de Starbucks que se parece al café), y bueno por la tarde también y a cualquier hora realmente…Gracias por tu simpatía y por tener paciencia con estos europeos que no tienen sentido del pudor y no se quieren casar (aunque ya ves que siempre hay esperanza!). Emilie, ha sido un placer compartir contigo este tiempo en el labo, desde que llegaste aprendí mucho sobre la diferencia entre la Francia de París y la Francia de Toulouse, me falta entender cual es más francés pero me temo que no hay respuesta. Patri, gracias por tu entusiasmo, me encanta como unas cosas pequeñas te puedan dar tanta alegría, de tu attitud deberíamos aprender todos! Raski, me faltan unos genotipos de ratones BHT que pero se genotipan como BHY. Es broma! Gracias por todo tu trabajo y por llamarme cuando no me entero de que la gente ha bajado a comer. Si no fuera por ti comería muchas más veces en frente del ordenador y esto no es bueno para un animal social como el hombre. Chiara, mi compatriota, gracias por ser siempre tan colaborativa, por cuidar de mis ratones cuando no estoy, pero sobretodo por traer tarta los días que más me apetece (basicamente todos). Raquel (R2) gracias por preocuparte por mi, tu sentido maternal se nota hasta en el labo, Enrique es muy afortunado y por esto es la monería de niño que es, más rico imposible. Fabio, el riminese que vino a jugar a voleyplaya en Madrid…aunque no te quedaste mucho tiempo en el labo no puedo no agradecerte por la alegría que llevaste…como nos pudimos reir contigo! Tus nuevos compañeros son muy afortunados! Gracias Lechuga por todo lo que haces en el labo, por trabajar tanto y siempre sonriendo y contando las noticias más curiosas que nadie sabe de donde las sacas pero oye, que también es cultura. Mil millones de gracias por ayudarme durante un año entero a poner a punto el sorting de pulmón y por recogerme del suelo cada vez que me he desmayado (en Salamanca no porque no estabas). Gracias también a Marta por intentar poner algo de orden en el labo, ya se que no te lo ponemos facil. Gracias a Carmen y David, las columnas de Oncología Experimental, por ser mi guía cada vez que lo necesitaba. Esta tesis es también el resultado de vuestras correcciones, vuestras sugerencias y vuestro tiempo y os lo agradezco mucho.

Eva, Marta, Jareck C, Jareck S, Aga, Silvia, Matt, Miljiana, Kerstin, mis compañeros de La Caixa, empezamos juntos este largo recorrido doctoral, yo me voy la ultima y me llevo muy buen recuerdo de todos vosotros, sois especiales y no tengo duda de que vais a

tener mucho éxito en vuestras vidad cientificas y personales. Aga, a ti en especial te agradezco tu alegría y tu disponibilidad, tu energía contagiosa, tu ser tan buena persona y amiga…eres única y te quiero mucho! Evita, tu eres de lo mejor que me llevo de esta experiencia en España y en el CNIO. Una amiga siempre presente, disponible, paciente, cariñosa…necesitaba una persona como tu y te agradezco por estar ahí. Espero que la vida y las carreras nunca nos separen, pero ya se que no pasará, pues tu eres de estas personas especiales que saben como cuidar de una amistad como de una planta que necesita ser regada de vez en cuando, y de mimos, y de amor.

Un gracias muy grande también a todo el personal de las unidades del CNIO con las que he tenido el placer de trabajar. Mi trabajo no hubiera sido posible sin el vuestro. Gracias a Marta Cañamero y las chicas de Patolgía Comparada, así como a la Unidad de Citometría; a Mayte, Conchy, Vicky, Maribel, Isabel, por cuidar de mis bichitos. Miles de gracias a Diego, Ximo y Manu, el gran equípo de Confocal, durante los primeros dos años de mi tesis he pasado más tiempo en vuestros cuartos oscuros que en mi labo. Aprendí muchísimo y gracias a vuestro buen rollo soy fan total de la microscopía confocal. Gracias a Arancha por aguantar conmigo (y con una sonrisa) la frustración de sortear este 0.01% de células YFP positivas por las que tanto luchamos. Me alegro de que hayas encontrado un sitio donde valoren lo buena que eres en tu trabajo, te lo mereces. Y como que no, gracias a Celeste, por tener el valor necesario para meter mano en mi escritorio, y por anunciar la llegada de las 18:00 con tu sonrisa y tu alegría.

Finora ho ringraziato tutti coloro che hanno reso possibile questa intensa esperienza a Madrid e al CNIO, attraverso il lavoro gomito a gomito e il contatto giornaliero. Ma credo che sarebbe un errore non ricordarmi di chi invece ha dovuto sopportare la distanza e l'assenza. Anche il vostro appoggio é stato fondamentale. Grazie a mio padre, mio fratello e tutta la mia famiglia in Italia (compresa quella acquisita). Grazie alle mie amiche, Arianna, Claudia, Laura, che ogni volta mi hanno riaccolta come se non me ne fossi mai andata. Una squadra é sempre una squadra, anche quando i membri prendono altre strade, e in questi giorni piú che mai me lo state dimostrando, che se lancio una palla, anche da qui, voi siete lí a prenderla.

Grazie all'angelo che continua a guidarmi da un posto che io non conosco, spero di non averti delusa e di non farlo mai.

E per ultimo il piú importante, e il piú difficile da ringraziare, perché non trovo le parole per descrivere quanto siano stati importanti per me il tuo appoggio, la tua presenza, il tuo amore. Guido, vita mia, questo lavoro é stato fatto a quattro mani, questi risultati sono tanto miei quanto tuoi. Per avermi spinto a inseguire il mio sogno, per avermi seguito, per avermi aspettato, per esserti occupato di me quando il lavoro assorbiva tutta me stessa, per avere avuto pazienza, ma tanta pazienza, tonnellate di pazienza inimmaginabili per un solo essere umano! Per tutti i finesettimana passati con me in lab, per le tazze di té mentre preparavo i labmeeting, per aver imparato a usare "overnight" nelle conversazioni domestiche. Ho scelto il miglior compagno per la vita e non posso che ritenermi una donna incredibilmente fortunata. Grazie, amore mio.

1 SUMMARY

1 SUMMARY

We have investigated the first steps of K-Ras^{G12V}-induced transformation in the K-Ras^{LSLG12Vgeo} model of lung adenocarcinoma. One one side, we have followed expansion of the different lung cell types after K-Ras^{G12V} –activation in order to identify lung cancer initiating cells. We have found that alveolar type II cells are the most likely to initiate lung cancer when the oncogene activation occurrs in an unbiased fashion. On the contrary, preferential expression of K-Ras^{G12V} in the bronchiolar epithelium via adeno-Cre intratracheal infection or activation of the oncogene during lung embryonic development, leads to hyperproliferation of Clara cells. On the other side, we obtained the gene expression profile of the cells that expand following oncogenic K-Ras activation at very early time points. Thus, we have identified the Ets transcription factor Etv4 as a key mediator of K-Ras^{G12V}-induced transformation.

1 RESUMEN

Hemos estudiado los primeros eventos que acompañan la transformación inducida por el oncogen K-Ras en el modelo de adenocarcinoma pulmonar K-Ras^{LSLG12Vgeo}. Por un lado, hemos seguido la expansión de los distintos tipos celulares del pulmón tras la activación del oncogen K-Ras con el fin de identificar las células iniciadoras del cancer de pulmón. Hemos descubierto que las células alveolares de tipo II son las que con más probabilidad dan origen a los adenocarcinomas de pulmón cuando la expresión del oncogen se produce de forma aleatoria. Sin embargo, la activación preferential de K-Ras^{G12V} en el epitelio bronquiolar a través de la infección intratraqueal con el virus adeno-Cre, así como la expresión del oncogen durante el desarrollo embrionario del pulmón, producen una hiperproliferación de las células Clara. Por otro lado, hemos obtenido el perfil de expresión genica de la población celular que se expande tras la expresión de K-Ras^{G12V} a tiempos muy tempranos. Así, hemos podido identificar el factor de transcripción de la familia Ets, Etv4, como un mediador muy importante de la transformación inducida por K-Ras^{G12V}.

INDEX

2 ABBREVIATIONS

2 ABBREVIATIONS

adeno-Cre:	**Adeno**virus carrying the **Cre** recombinase gene
adeno-FlpO:	**Adeno**virus carrying the **O**ptimized **Fl**i**p**ase recombinase gene.
ALK:	**A**naplastic **L**ymphoma **K**inase
Aqp5:	**Aq**ua**p**orin **5**
ATI:	**A**lveolar **T**ype **I** cells
ATII:	**A**lveolar **T**ype **II** cells
BADJ:	**B**ronchio**a**lveolar **D**uct **J**unction
BASC:	**B**ronchio**a**lveolar **S**tem **C**ell
β-gal:	**β-gal**actosidase
β-geo:	**β-g**alactosidase / n**eo**micine
CC:	**C**lara **C**ells
CC10:	**C**lara **C**ell protein **10**
CCSP:	**C**lara **C**ell **S**ecretory **P**rotein
CIC:	**C**ancer **I**nitiating **C**ell
CSC:	**C**ancer **S**tem **C**ell
DAPI:	4',6-Diamidino-2-Phenylindole
EGFR:	**E**pidermal **G**rowth **F**actor **R**eceptor
Erk:	**E**xtracellular signal-**r**egulated **k**inases
ETS:	**E**-**t**wenty **s**ix DNA binding domain
ETV:	**E**ts **V**ariant **4**
GAP:	**G**TPase-**A**ctivating **P**rotein
GDP:	**G**uanosine **d**iphos**p**hate
GEF:	**G**uanine nucleotide **E**xchange-**F**actor
GF:	**G**rowth **f**actor
CGRP:	**C**alcitonin **G**ene-**R**elated **P**eptide

GTP: **G**uanosine **t**riphos**p**hate

HE: **H**ematoxilin & **E**osin

IF: **I**mmuno**f**luorescence

IHC: **I**mmuno**h**isto**c**hemistry

IRES: **I**nternal **R**ibosomal **E**ntry **S**ite

KI: **K**nock **I**n

KO: **K**nock **O**ut

LOH: **L**oss **O**f **H**eterozigosity

Mapk: **M**itogen-**a**ctivated **p**rotein **k**inases

Mek: **M**ap **E**rk **K**inase

NE: **N**euro**e**ndocrine cells

NFR: **N**uclear **F**ast **R**ed

NF1: **N**euro**f**ibromatosis **1**

NSCLC: **N**on-**S**mall-**C**ell **L**ung **C**ancer

PCR: **P**olymerase **C**hain **R**eaction

PDGFR: **P**latelet-**D**erived **G**rowth **F**actor **R**eceptor

PDK1: **P**hosphatidylinositol-**D**ependent **K**inase-**1**

PFU: **P**laque-**f**orming **u**nit

PI3K: **P**hosphoinositol-**3**-**K**inase.Fosfoinositol-3-kinase

PKB: **P**rotein **K**inase **B**

PKC: **P**rotein **K**inase **C**

RB: **R**etino**b**lastoma

RT: **R**oom **T**emperature

RT-PCR **R**etro**t**rasncription of the RNA followed by **PCR**

RTK: **R**eceptor **T**yrosine **K**inase

SCLC: **S**mall-**C**ell **L**ung **C**ancer

Scgb1:	Secretoglobin 1
SPC/SfpC:	**S**urfactant **P**rotein **C**
TF:	**T**ranscription **F**actor
Tg:	**T**rans**g**ene
WT:	**W**ild **T**ype
X-Gal:	β-galactosidase enzymatic reaction
YFP:	**Y**ellow **F**luorescent **P**rotein
4-OHT:	4-hidroxitamoxifen

3 INTRODUCTION

3.1 LUNG CANCER

3.1.1 Statystics of lung cancer

Lung cancer is the leading cause of cancer deaths worldwide, with around 1.1 million deaths per year, accounting for 17.8% of the total deaths caused by cancer. The frequency of lung cancer is slightly higher in the west countries (22%) as compared to the less-developed countries (14%) (Travis et al., 2004). Also, this type of cancer is 2.7 times more frequent in males than in females (Travis et al., 2004). Nevertheless, in some occidental countries, like United States, where historically a higher number of women have been exposed to tobacco smoke, the percentage of diagnosed lung cancers and lung cancer-related deaths is very similar between the two genders (Siegel et al., 2013) **(Figure 1)**. Overall, smoking is the most important risk factor for lung cancer, as 85-90% of the cases can be attributed to tobacco smoking in men (Travis et al., 2004). This type of cancer has a very bad prognosis, with only a 10% 5 years survival in the majority of the countries.

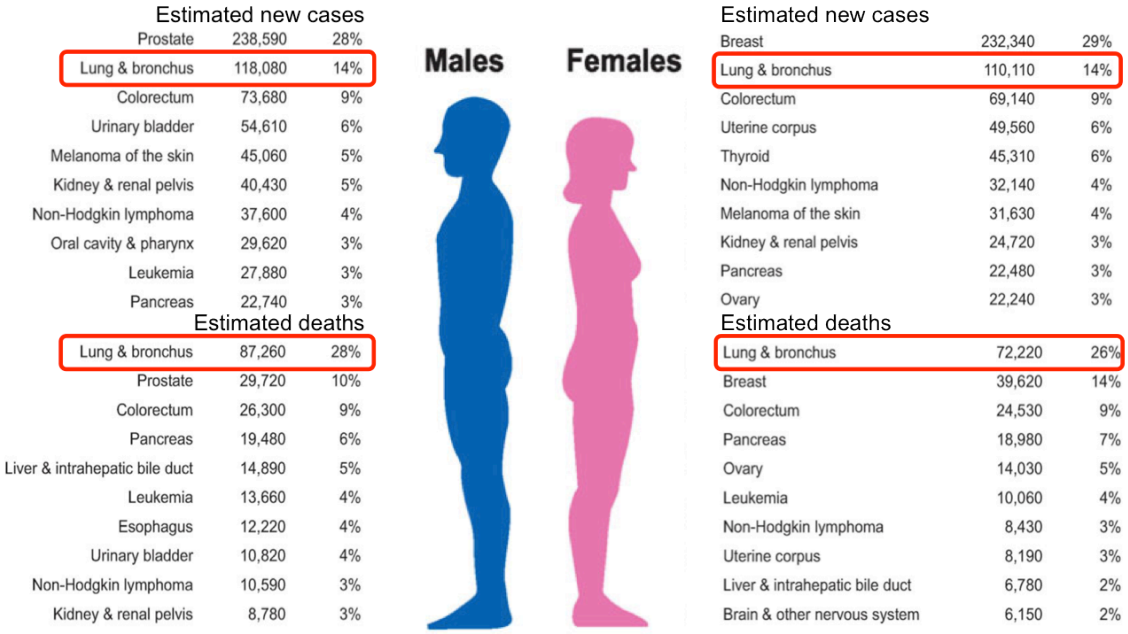


Figure 1: estimated percentage of new cases of cancer and cancer-related deaths in the United States in 2013 (adapted from Siegel et al. 2013).

3.1.2 Subtypes of lung cancer

Lung cancer can be classified into two big groups, according to histopathological criteria: Small Cell Lung Cancer (SCLC) and Non Small Cell Lung Cancer (NSCLC) **(Figure2)**. The first one is characterized by the presence of small cells that barely have room for any cytoplasm. They can be round, oval or spindle-shaped, with small or no nucleolus and they characteristically express neuroendocrine markers (Travis et al., 2004). This subtype accounts for around 18% of all lung cancer cases and is very poorly responsive to chemio and radiotherapy (Meuwissen and Berns, 2005). The 5 years survival rate is only 5% (Worden and Kalemkerian, 2000). The cases that do not correspond to these features are classified as Non Small Cell Lung Cancers, and they account for 80% of all the malignant lung cancers **(Figure 2)**. NSCLC can be further divided into 3 main subgroups according to histopathological criteria: adenocarcinoma, squamous cell carcinoma, and large cell carcinoma (Travis, 2002) **(Figure 2)**. The first two subgroups are the most frequent ones inside the NSCLC subgroup. (Meuwissen and Berns, 2005). The 5 years survival rate of these cancers is around 14% (Travis et al., 1995). Squamous cell carcinoma histologically displays the characteristics of squamous epithelium, while large cell carcinoma is composed by big, mostly undifferentiated, cells.

The work reported here is particularly focused on lung adenocarcinoma, which is the most frequent type of NSCLC (40%) **(Figure 2)**. Lung adenocarcinoma is an epithelial cancer with a glandular aspect, which can assume different architectural patterns: acinar, papillary, bronchioalveolar, solid with mucin production or a mixture of different patterns (Travis et al., 2004).

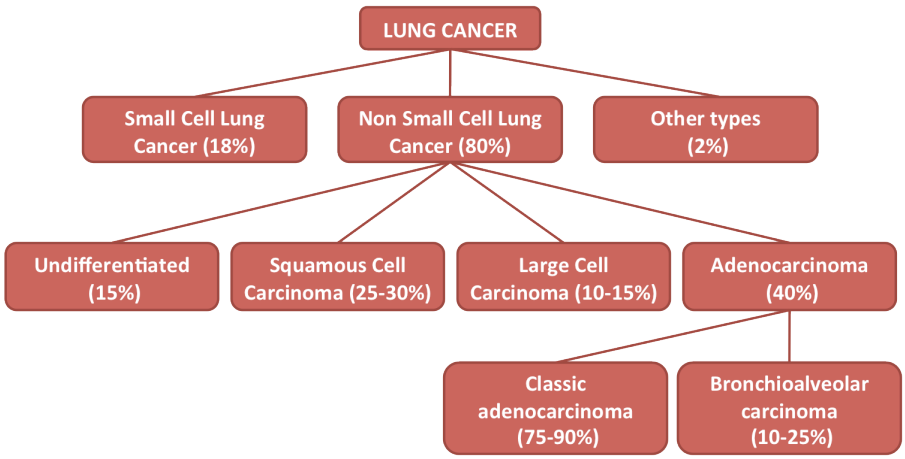


Figure 2: classification of lung cancer and relative frequence of the different subtypes. Data from Travis et al., 2002.

3.1.3 Frequent mutations in lung cancer

Specific genes are frequently found mutated in lung cancer, for example inactivating mutations of the TP53 tumor suppressor gene are found in 50% of NSCLC and 70% of SCLC. The second most frequent mutations are the ones involving the retinoblastoma protein (RB) signaling pathway. In fact, loss of RB expression occurs in 80-100% of advanced SCLC, although it is less frequent (15%) in NSCLC (Travis et al., 2004). Nevertheless, there are other mutations that, although not very frequent in the overall lung cancer cases, are the predominant ones in a given histopathological subtype. For example, mutations in the codon 12 of the K-RAS locus have been found in approximately 30% of lung adenocarcinomas, even though this mutation is extremely rare in other types of NSCLC as well as in SCLC (Travis et al., 2004). K-RAS mutation is also a marker for bad prognosis in NSCLC (Graziano et al., 1999). While K-RAS mutation is predominant in adenocarcinoma patients with a smoking history, mutations in other genes such as the tyrosine kinase receptor EGFR or ALK (anaplastic lymphoma kinase) are more frequent in never smokers (Pao et al., 2004a); (Soda et al., 2007) **(Figure 3)**. EGFR, ALK and K-RAS mutations seem to be mutually exclusive, as they are very rarely found altered together in the same tumor, supporting the hypothesis of an oncogene-induced toxicity (Gerber and Minna, 2010). Importantly, mutations of EGFR confer sensitivity to the tyrosine kinase inhibitors gefitinib and erlotininb (Lynch et al., 2004); (Pao et al., 2004b); (Paez et al., 2004).

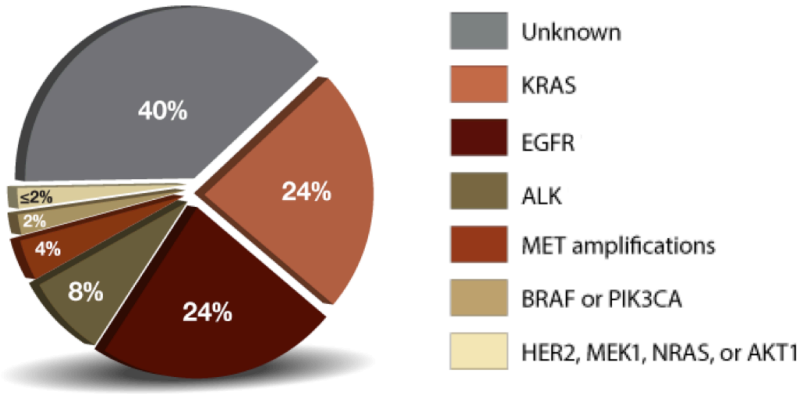


Figure 3: presence of single driver mutations in NSCLC. (Adapted from Kris et al., 2011).

3.2 THE RAS PROTEINS

3.2.1 The Ras family

The Ras family is composed of four members: H-Ras, N-Ras, K-Ras4A and K-Ras4B, the last two resulting from alternative splicing of two different exons in the K-Ras locus (Capon et al., 1983); (George et al., 1985); (McGrath et al., 1983); (Shimizu et al., 1983). K-Ras4B is the more abundant of the two isoforms, and therefore it is generally referred to simply as K-Ras (Pells et al., 1997); (Wang and You, 2001). The Ras proteins are characherized by a highly conserved amino-terminal domain (between 90 and 100% of homology). On the contrary, the carboxy-terminal domain contains the so-called hypervariable region. Here the -CAAX domain was identified, as the one responsible for Ras anchorage to the plasma membrane. All the Ras proteins are post-translationally modified in this domain by the cytosolic enzyme farnesyl-transferase, which acts by adding a farnesyl group. This modification drives all the Ras proteins to the endoplasmic reticulum, where the tripeptide –AAX is eliminated and the proteins are methylated. From this point on, the different Ras proteins undergo different post-translational modifications. K-Ras goes to the plasma membrane and interacts with its negative charges thanks to its polylisine domain. H and N-Ras undergo another lipidic modification, namely palmitoylation, by the palmitoyl-transferase enzyme (two palmitoylations for H-Ras and one for N-Ras). These differences in post-translational modifications are responsible for the different sublocalizations to the membrane and different signaling pathways activated. So, H-Ras is localized to the lipid rafts, very rigid regions of the plasma membrane, while K-Ras is located in more disorganized membrane regions (non lipid rafts, which are much more fluid and allow higher mobility) (Hancock, 2003).

3.2.2 Ras proteins activation

Ras proteins are GTPases, meaning that they have the intrinsic capacity to hydrolyze GTP into GDP. Ras is in its active state when GTP-bound, and in this conformation it acts as a cytoplasmic mediator of signals coming through the membrane. This active state is

transient, as GTP hydrolysis leads to K-Ras inactivation. Other proteins take part in the regulation of this Ras cycle **(Figure 4)**. Namely, Ras activation is catalyzed by GEFs proteins (guanine nucleotide exchange factors), which help GTP binding to Ras, and GDP release. The first GEF described was Cdc25 in *Saccaromyces cerevisiae* (not to be confused with the cell cycle-related protein) (Broek et al., 1987); (Robinson et al., 1987). Its mammalian equivalent, SOS, was lately identified and cloned. (Downward et al., 1990); (Shou et al., 1992); (Wei et al., 1992); (West et al., 1990); Wolfman y Macara, 1990). Other proteins that act as SOS have been also identified **(Figure 4)**. The Grb2 protein acts as a connection between SOS and signals that come from the membrane. In fact Grb2, thanks to its SH2 domains, interacts with the phosphorilated, active, Tirosine Kinase Receptors (RTKs) (Karnoub and Weinberg, 2008); (Malumbres and Barbacid, 2003). Grb2 is in turn controlled by other proteins including the ACK and Sprouty families, resulting in impeded SOS activation (Pao-Chun et al., 2009); (Hanafusa et al., 2002) **(Figure 4)**.

As any other biological process, the transmission of a signal needs to be controlled in time. As the intrinsic Ras GTPase activity is quite weak, other proteins called GAPs (GTPase-Activating Proteins) help to catalyze the GTP hydrolysis. The first RasGAP was identified based on the observation that some soluble factors were able to increase the activity of wild type H-Ras and N-Ras, but not of their oncogenic forms (Gibbs et al., 1988); (Trahey and McCormick, 1987); (Trahey et al., 1988); (Vogel et al., 1988). Moreover, this demonstrated a clear connection with the RTKs, as PDGFR activation was able to promote RasGAP phosphorylation, thus impairing its translocation to the plasma membrane (Molloy et al., 1989). Interestingly, the importance of those RasGAPs was not fully evident until a second RasGAP, NF1, was found altered in patients suffering from neurofibromatosis (Ballester et al., 1990); (Martin et al., 1990); (Wallace et al., 1990); (Xu et al., 1990). Other GAP proteins have lately been identified **(Figure 4).**

Alterations in Ras activity have been described in many human cancers. Notably, the mutations in codon 12 of the K-RAS locus identified in lung adenocarcinoma leads to the disruption of the K-RAS GTPase activity. In this way, K-RAS constantly remains in its GTP-

bound status and it is thus constitutively active. Mutations in codons 13 and 61 have the same effect, although they are less frequent (Karnoub and Weinberg, 2008); (Malumbres and Barbacid, 2003).

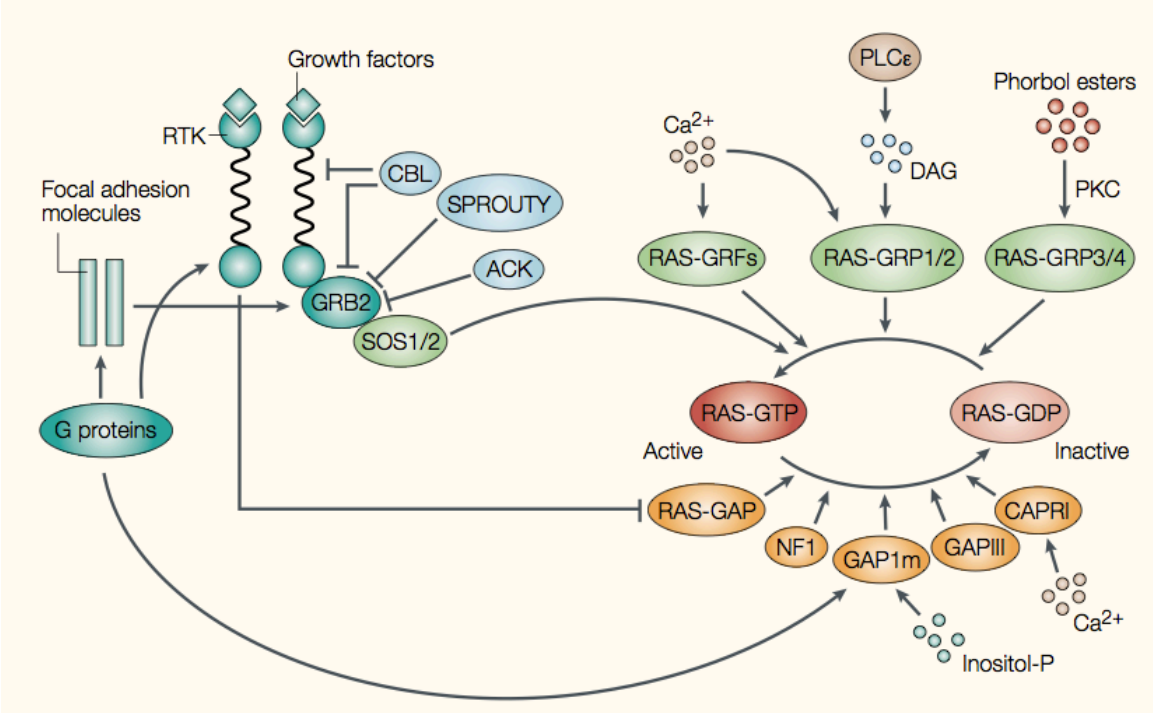


Figure 4: Ras regulation (adapted from Malumbres and Barbacid, 2003). Ras is activated following distinct extracellular stimuli, mediated by the RTKs, as well as G proteins, adhesion molecules and second messengers. Those stimuli activate various RasGEFs (in green). Inhibitors of RTK-mediated Ras activation include the Sprouty and Ack protein families, which bind to Grb2 and impede the activation of SOS. Cbl has got a similar function, but it can also block signal transduction through RTKs ubiquitination. Several RasGAPs (in orange) also downregulate Ras by catalyzing the GTP hydrolysis. DAG=diacylglycerol, PLC=phospholipase C.

3.2.3 The Ras protein effectors

Ras effectors are defined as proteins that are activated in a Ras-dependent manner, as a consequence of stimulation by either growth factors (GFs) or hormones. The effectors are generally localized at the plasma membrane, where they are preferentially bound to the Ras active form, Ras-GTP (Downward, 2003); (Malumbres and Barbacid, 2003) **(Figure 5)**.

The first Ras effector described was the serine/threonine kinase Raf. The Raf family consists of three proteins, C-Raf, B-Raf and A-Raf. When Raf proteins are active, they

have the capacity to phosphorilate Mek1 and Mek2 (Wellbrock et al., 2004). The activation of these two proteins leads to the phosphorylation and subsequent activation of Erk1 and Erk2. Once Erk is activated it translocates to the nucleolus where it can bind a wide range of nuclear factors, including the ETS family of transcription factors. The best known of these is probably Elk, a protein that regulates the expression of genes such as Fos or c-Jun. As a consequence of the stimulation of this transcription factors activity, specific cell cycle regulatory proteins are synthesized. Among these the D-type cyclins, that are crucial in cell cycle initiation and progression through G1 phase (Marshall, 1999). In summary, the activation of Raf promotes cell cycle progression due to the activation of the Raf/Mek/Erk pathway, commonly known as the mitogenic or proliferative Ras pathway **(Figure 5)**.

A second cascade activated by Ras is the phosphatidylinositol 3-kinase (PI3K) pathway **(Figure 5)**. The activation of PI3 kinases produces lipid-based second messengers like phosphatidylinositol 3,4,5 triphosphate. These lipid factors can then activate a wide number of enzymes such as Pdk1. This last is important for the activation of the AGC family of kinase proteins, which includes Akt, PKC and p700S6K. The last two proteins are specially implicated in signal transduction. Akt, also known as Pkb, has a key anti-apoptotic role. For this reason the PI3K/Pdk1/Akt pathway is also known as the Ras cell survival pathway. Besides, Akt is also involved in cell mobility through cytoskeletal proteins activation, such as RhoB. The PI3K proteins also lead to the activation of Rac, a member of the Rho family that not only has a role in cytoskeleton regulation but also in the stimulation of specific transcription factors such as Nf-kB (Nuclear Factor kB) (Krasilnikov, 2000).

The Ral-GEFs are a third class of Ras effectors **(Figure 5)**. The Ral-GEFs are nucleotide-exchanging factors of the Ral family (small size GTPase proteins). Ral drives the inactivation of Cdc42 and Rad, two proteins that work together in the cytoskeleton regulation. On the other hand, Ral activation inhibits the transcription factor Forhead, of the FoxO family, leading to expression of the cell cycle inhibitor $p27^{kip1}$ (Essers et al., 2004).

In summary, Ras proteins play a universal role in cellular signaling transduction, due to the wide range of effectors they are able to activate, that brings to an effect in most cellular biological processes (Malumbres and Barbacid, 2003) **(Figure 5)**

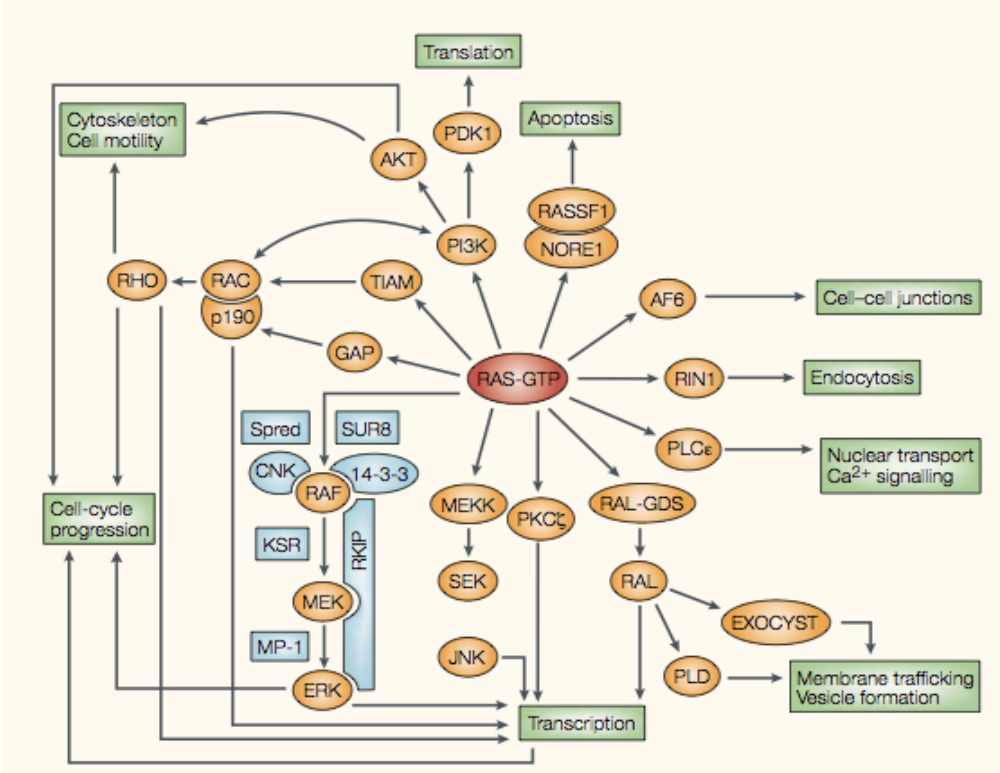


Figure 5: RAS effectors and their corresponding biological responses (adapted from Malumbres and Barbacid 2003). Active RAS-GTP induces a wide variety of cellular processes, such as transcription, translation, cell-cycle progression, apoptosis or cell survival, through direct interaction with various effectors. GAP proteins also interact with RAS-GTP and might also act as effectors. Modulators of some of these pathways are also indicated. The blue boxes represent adaptor complexes.

3.3 THE USE OF MOUSE MODELS FOR THE STUDY OF LUNG CANCER.

It is known that spontaneous tumors that arise in mice are similar in morphology, histopathology and molecular alterations to human cancers (Meuwissen and Berns, 2005). Thus, appropriate models can convert mice into a very useful tool to better understand cancer pathology, to discover tumor markers, as well as to develop and test new antitumoral strategies.

The first mouse models for lung cancer were obtained based on the observation that certain inbred strains were more susceptible to the development of spontaneous

tumors (Meuwissen and Berns, 2005). But obviously, although those models recapitulated the spontaneous occurrence of tumors in humans, the experimental conditions couldn't be properly controlled and lack of reproducibility was a big problem. Later it was observed that those lines were also more susceptible to carcinogenic chemical agents, and that the results obtained with carcinogens were highly reproducible (Malkinson, 1989); (Shimkin and Stoner, 1975). The historically used carcinogens were tobacco-derived nitrosamines, polycyclic aromatic hydrocarbons and ethyl carbamate. In those experiments, following the carcinogen application, foci of hyperplasia appeared. Some of them eventually developed into adenomas, and a certain number also became malignant, giving rise to adenocarcinomas with local invasion (but no metastasis). Those adenocarcinomas were positive for SPC (Surfactant Protein C, also called Sfpc), a marker for Alveolar Type II cells (ATII), a characteristic that is also recurrent in human lung cancer (Meuwissen and Berns, 2005) (for a description of lung cell types see section 3.4.1 and **Figure 8**). Further studies on these tumors indicated that K-Ras was mutated already at the hyperplastic stage (Horio et al., 1996), pointing at K-Ras mutation as the driving alteration for tumor initiation. c-Myc was also found frequently overexpressed (Re et al., 1992). In the most advanced stages, inactivating mutations of tumor suppressor genes like Trp53, Rb or Apc were also frequently detected (Malkinson, 2001).

Nevertheless, the random mutagenesis process of those first mouse models made it difficult to control the experimental conditions. For this reason, more sophisticated models have been developed in the last decades, like transgenic (Tg) mice that express oncogenes frequently mutated in human and mouse spontaneous tumors, or knock out (KO) mice in which tumor suppressor genes are inactivated. Up to date, a high number of models for different oncogenes and tumor suppressor genes have been developed (Meuwissen and Berns, 2005); (Kwon and Berns, 2013). Nevertheless, in the first genetically engineered mouse models, the oncogene expression or tumor suppressor removal occurred in a constitutive and ubiquitous fashion. In this aspect, those models were still quite far from the real situation in which a tumor develops from a sporadic

mutation occurring in a specific organ (Jonkers and Berns, 2002). Moreover in many occasions the tumor was induced by non-physiological overexpression of oncogenes, due to the use of very strong promoters in the transgene constructions. The problem of those germline transgenic and knock out mice was partially solved by the use of conditional alleles, which better mimic the sporadic tumorigenesis process, allowing the control of the organ and the temporal window in which the mutated gene is expressed. The best example of such an advance is the Cre/loxP recombination system (Gu et al., 1993); (Kuhn et al., 1995). In the case of tumor suppressor genes, a determined region of the gene of interest is flanked by two loxP sequences. The expression of the Cre recombinase in a given tissue and/or at a given moment allows the controlled elimination of the target protein expression. In case that the gene of interest is an oncogene, a transcriptional STOP cassette is inserted at the 5'end, flanked by two loxP sites. In this way, Cre expression is able to eliminate the STOP cassette, thus allowing the expression of the oncogene in the desired organ and at the desired time point. A further evolution of this conditional expression system is the knock in (KI) approach. In this case, the expression of the oncogene of interest is under the control of its endogenous promoter, thus avoiding the non-physiological overexpression produced in the transgenic models by the use of highly active promoters. Thanks to the Cre/loxP system better mouse models for lung cancer started to be generated. For example, elimination of Tp53 and Rb (genes which are frequently found mutated in SCLC) in mouse lung cells, allowed the development of a SCLC mouse model (Meuwissen et al., 2003). In summary, the tools available today allow us to conditionally activate or inactivate specific genes in distinct cell types at any desired moment. Yet, the major drawback of current mouse model systems that combine multiple genetic modifications is the time required to bring together all the desired mutations in a single genome. Nevertheless some groups recently started to develop strategies aimed at speeding up this process (Dow et al., 2012). For example it is now possible to replace crossbreeding by consecutive genetic modification of embryonic stem cells (ES) (Zhou et al., 2010) derived from strains that already harbor an important number of the mutations required to complete the model

(Huijbers et al., 2011). Furthermore, we can now substitute genetic modifications in ES cells partly by somatic gene transfer using an array of specifically designed adenoviral and lentiviral vectors (Xia et al., 2012). Thanks to these advances we can introduce a number of tumor-specific driver mutations into the appropriate target cell in order to reproduce most of the unique features of that particular tumor type, allowing us to answer the question of how individual mutations can contribute to the tumor phenotype and how they influence intervention strategies. Also, we become able to combine activation of oncogenes with tumor suppressors inactivation to produce more aggressive tumor models, as well as to delete single components of the downstream pathways in order to assess tumor addiction to defined effectors.

3.3.1 Mouse models for NSCLC.

One of the first models for lung adenocarcinoma was developed in the lab of Harold Varmus (Fisher et al., 2001). In their transgenic CCSP-rtTA; ($tetO_7$) CMV-K-Ras^{G12D} mice, in the presence of doxycycline, expression of Cre recombinase was supposed to occur only in the bronchiolar Clara cells (due to the use of the CCSP, Clara Cell Secretory Protein, promoter). Nevertheless they obtained Cre expression also in the majority of ATII cells. In any case, the advantage of this TetOn system was that upon doxycycline withdraw the oncogene was no longer expressed and the adenocarcinomas previously formed started to regress. In parallel, another model based on the K-Ras^{G12V} mutation was generated in the lab of Anton Berns (Meuwissen et al., 2001). In this case the oncogene sequence was preceded by the GFP gene, flanked by loxP sites, allowing GFP expression prior to adeno-Cre infection. But also in this case the system was not very physiological, as the oncogene was transcribed under the control of the β-actin promoter. To this respect, a substantial advance came from the lab of in Tyler Jacks, who developed a KI model characterized by a latent oncogenic allele, K-Ras^{G12D}, expressed from its endogenous promoter following a spontaneous recombination event (Johnson et al., 2001). Lately, the same lab developed a more sophisticated version of this model in order to avoid the spontaneous random recombination step. In this case, the expression of a resident K-Ras^{G12D} oncogene was prevented by the insertion of a

STOP cassette (Johnson et al., 2001). This cassette was flanked by loxP sites, in a way that infection of the animals with adenovirus carrying the Cre recombinase gene (adeno-Cre), led to the elimination of the STOP sequence and transcription of the oncogene. Noteworthy, in both Anton Berns and Tyler Jacks's models, hyperplasias and adenomas appeared 4 weeks after adeno-Cre infection, and adenocarcinomas were also detectable after 9-12 weeks (Jackson et al., 2001); (Meuwissen et al., 2001). Nevertheless, no metastases were observed, but only local invasion.

In our lab, another KI model was generated, in which the K-Ras^{G12V} mutation was introduced in the endogenous K-Ras locus, and a STOP cassette flanked by loxP sites was placed in the 5'end in order to prevent its expression (Guerra et al., 2003) **(Figure 6)**. An additional advantage of this model was the presence of a β-Geo cassette introduced in the 3'end of the K-Ras locus, preceded by an IRES sequence (Internal Ribosomal Entry Site) (Friedrich and Soriano, 1991). Through this bicistronic system the oncogene expression is accompanied by the parallel expression of the β-Galactosidase. In this way through a X-Gal reaction the K-Ras^{G12V} expressing cells can be identified already at the very first stages of tumor formation. In a manner similar to the previously described models, expression of the oncogene also leads to the development of adenomas and adenocarcinomas.

In order to have an alternative and more ubiquitous way of Cre recombinase expression, apart from adeno-Cre intratracheal infection in the lung, our K-Ras KI model was also crossed with the RERTn line developed in our laboratory by Victoria Campuzano. In this strain, a modified Cre recombinase is expressed under the promoter of the RNA polymerase II, thus providing its ubiquitous expression (Guerra et al., 2003). This Cre recombinase is merged to the binding domain of the steroid hormone estrogen receptor (ER), previously modified (ERT2) in a way that prevents its binding to the endogenous steroids. The addition of the synthetic steroid 4-hydroxy-tamoxifen (4-OHT) provokes the dissociation of the ERT2 from HSP40 that anchors it in the cytoplasm, resulting in the translocation of the Cre recombinase to the nucleus (CreERT2, (Brocard et al., 1997)). The administration of 4-OHT leads to postnatal expression of the K-Ras^{G12V} endogenous

oncogene in most of the mouse tissues. However most of the cell types tolerate the expression of endogenous physiological levels of the K-Ras^{G12V} oncogene without developing neoplastic lesions other than adenomas and adenocarcinomas in the lung (Guerra et al., 2003).

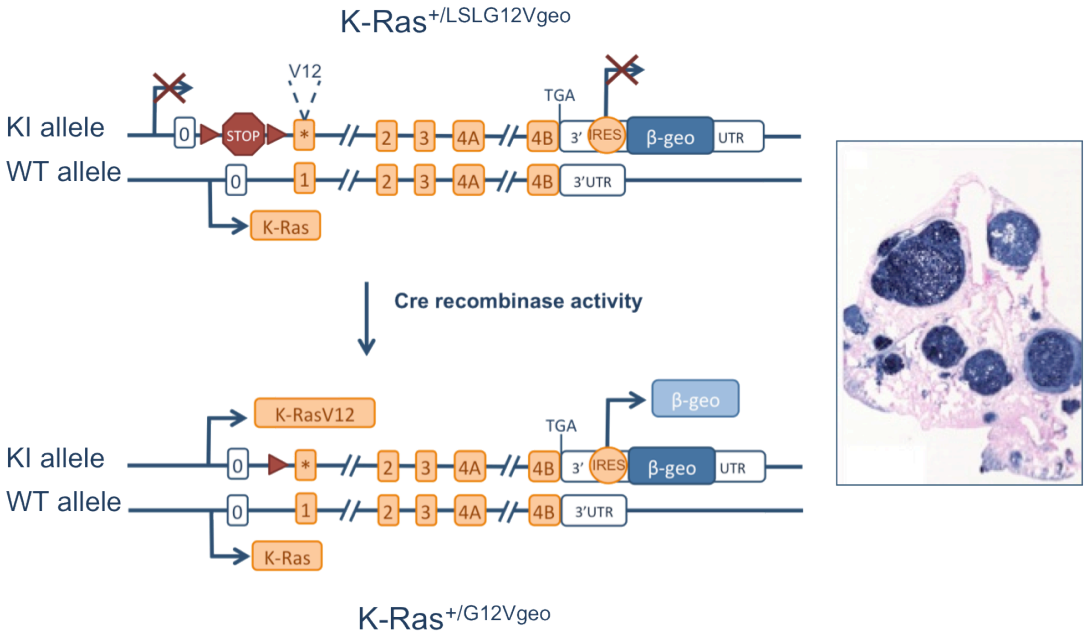


Figure 6: the K-Ras^{LSLG12Vgeo} model. Left panel, before activation of the Cre recombinase, the STOP cassette introduced in the K-Ras locus impairs the transcription of the bicistronic messenger. Following adeno-Cre infection the Cre recombinase recognizes the loxP sites, removes the STOP cassette and allows expression of K-Ras^{G12V} and of β-Geo in the same messenger RNA. The IRES sequence drives entry into the ribosomes and allows β-Gal translation. **Right panel**, section from a whole mount X-Gal stained lung from a K-Ras$^{+/LSLG12Vgeo}$ mouse six months after adeno-Cre inoculation. Adenomas and adenocarcinomas appear in blue, following expression of the β-Gal reporter.

In our lab a more aggressive lung adenocarcinoma model was also developed, based on a previous report describing that the murine K-Ras wild type (WT) allele displays tumor suppressor properties and, as a consequence, is frequently lost during lung tumorigenesis (Zhang et al., 2001). Indeed, this loss of heterozigosity (LOH) can be detected in about 50% of human lung cancer samples that carry K-Ras activating mutations (Li et al., 2003). In order to mimic spontaneous LOH a conditional K-Ras allele was developed in our lab (Puyol et al., 2010) by introducing two loxP sites flanking the exon 1 of the wild type K-Ras sequence This strain was crossed to the previously

described inducible K-Ras^{G12V} mouse line. In the resulting mice, designated as K-Ras$^{lox/LSLG12Vgeo}$, induction of the Cre recombinase results in the deletion of the wild type copy of K-Ras and the concomitant induction of the oncogenic mutation **(Figure 7)**. Elimination of the wild type allele and expression of the oncogenic K-Ras^{G12V} results in an earlier onset of tumors and a reduction of mice lifespan.

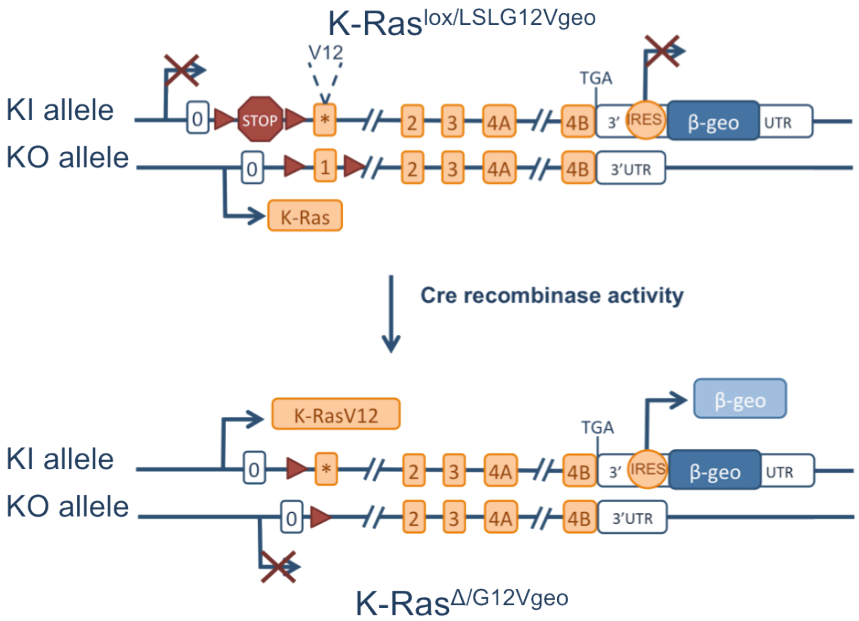


Figure 7: the K-Ras$^{lox/LSLG12Vgeo}$ model. Before activation of the Cre recombinase, the STOP cassette introduced in the K-Ras locus impairs the transcription of the bicistronic messenger. Following adeno-Cre infection the Cre recombinase recognizes both the loxP sites flanking the STOP cassette in the knock in allele and the ones flanking the exon1 of the wild type allele, thus allowing expression of K-Ras^{G12V} and of β-Geo from the targeted locus and impairing expression of a functional WT K-Ras from the second allele. Δ= KO K-Ras allele in which the exon 1 has been removed by the Cre recombinase.

Other mouse models for lung adenocarcinoma have been engineered based on mutations other than K-Ras, which have also been observed in lung cancer patients (Kwon and Berns, 2013). B-Raf^{V600E} inducible knock in mice show a rather benign tumor phenotype (Dankort et al., 2007). On the contrary, lung-specific switching on of the conditional EgfrL858R mutant allele in mice results in diffuse lung cancer resembling bronchioalveolar carcinomas (Regales et al., 2007). ALK (Anaplastic Limphoma Kinase) encodes a receptor tyrosine kinase that is not expressed in the normal lung. However,

the EML4-ALK fusion oncoprotein is often found in NSCLC patients and causes constitutive tyrosine kinase activity in the lung (Soda et al., 2007).

EML4-ALK represents about 5% of all NSCLC and is mutually exclusive with K-Ras and EGFR mutations (Gerber and Minna, 2010). Transgenic mice expressing EML4-ALK fusion protein develop multiple lung adenocarcinomas shortly after birth (Soda et al., 2008).

3.4 THE CELLULAR ORIGIN OF CANCER.

According to the classical clonal evolution model, cancer is considered as an evolutionary process that results from the accumulation of somatic mutations and other alterations, in the progeny of a normal cell, leading to a selective growth advantage of the mutated cells and ultimately to uncontrolled proliferation **(Figure 10)**. The tissues from which most cancers arise are maintained by multipotent and/or unipotent stem cells and their progeny, the committed progenitors or transient amplifying cells, that ultimately give rise to the non-dividing terminally differentiated cells that accomplish the various tissue functions. Thus, the genetic event leading to tumor initiation can occur at any level of this hierarchy, and the target of the initial oncogenic insult may dictate the tumor subtype arising in a given organ. Although many genes that lead to different types of cancer when mutated have been identified, the cells that respond to the driving oncogenic stimulus remained elusive for a long time. More recently, genetic lineage-tracing experiments that allow expression of oncogenes and/or deletion of tumor suppressor genes in defined cell lineages have helped to define the cellular origin of different solid tumors (Blanpain, 2013). In most of the models, oncogene expression in a given tissue leads to transformation of some cells but not others, although it remains unclear why some cell types are sensitive to oncogene-induced tumorigenesis while others are resistant. Understanding such mechanisms, and whether they are conserved across different tissues, will be important for the development of strategies to possibly block tumor progression.

3.4.1 Cell of origin of lung cancer.

The airways system is divided into trachea, bronchi and bronchioles, which branch and end in the alveoli, the site of gas exchange (Rock and Hogan, 2011). The trachea and major bronchi show the presence of a basal epithelium that primarily contain basal cells expressing cytokeratin 5 (CK5) and cytokeratin 14 (CK14). Ciliated cells expressing the Foxj1 marker and Clara cells expressing the CCSP/Scgb1/CC10 antigen compose the luminal epithelium **(Figures 8 and 9)**. The bronchioles are more distal, monolayer structures, made up of ciliated and Clara cells. Neuroendocrine (NE) cells, which characteristically express the calcitonin-gene related peptide (CGRP) marker, are scattered through the trachea, bronchi and bronchioles, although they are more frequent in the last ones **(Figure 8)**. The distal epithelium, or alveoli, is made up of Alveolar Type I (ATI) squamous cells that cover the alveoli, and ATII secretory cells. These two types of pneumocytes respectively express aquaporin 5 (AQP5) and surfactant protein C (SPC/Sfpc) **(Figure 8)**. In addition, the lung harbors many other cell types of non-lung origin such as macrophages, endothelial cells, pericytes and fibroblasts (Rock and Hogan, 2011).

Different stem cell populations have been identified as responsible for the maintenance and repair of the different airway epithelia. Trachea and bronchi harbor multipotent basal stem cells that can differentiate into several lineages including transit amplifying luminal progenitors and postmitotic basal and luminal cells (Rock et al., 2009). However, in the terminal bronchioles, Clara cells constitute a self-sustained population that contains long-term stem cells (Rawlins et al., 2009). On the contrary, the precise mechanisms maintaining and repairing alveoli are not well understood.

Since SCLCs express neuroendocrine markers, they are thought to arise from NE cells or NE progenitors. Mutations of p53 and Rb tumour suppressor genes have been found in the majority of human SCLCs (Beasley et al., 2003) and their inactivation in mice recapitulates the key features of human SCLC (Meuwissen et al., 2003).

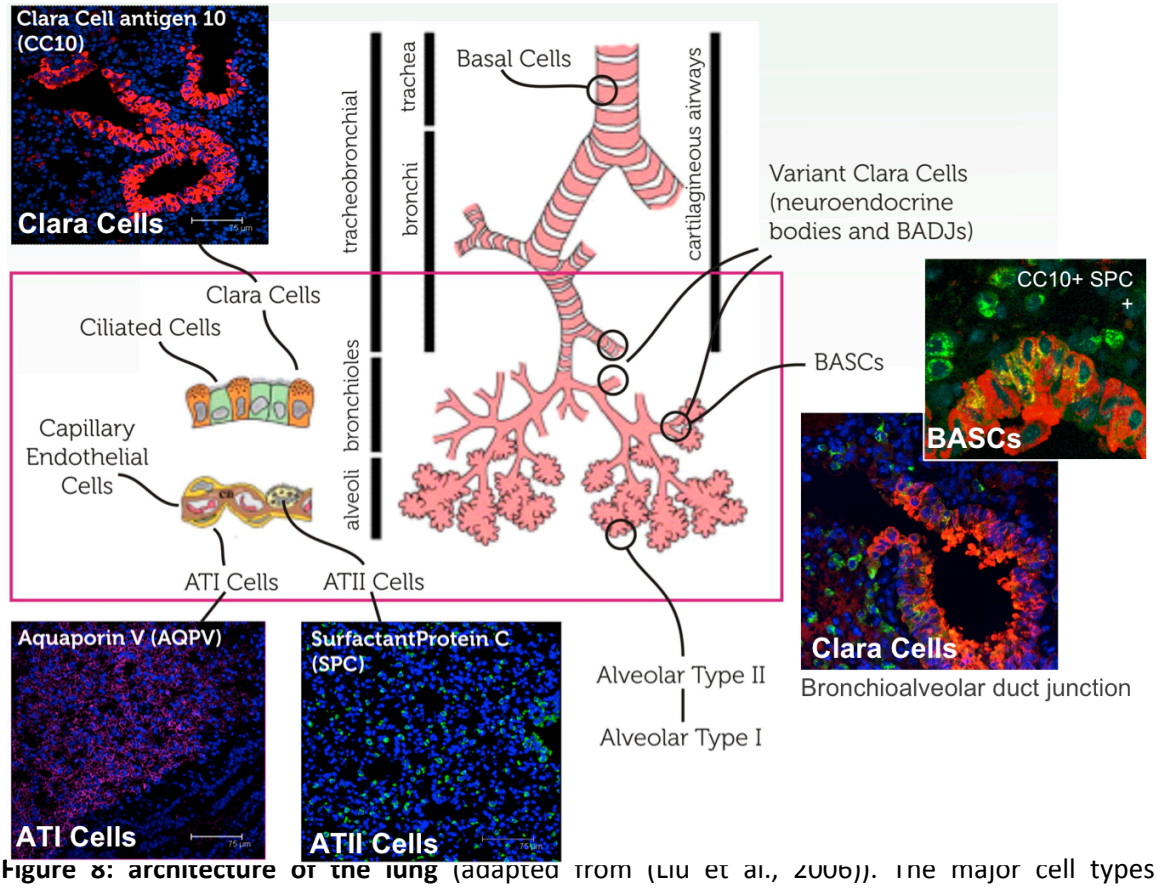


Figure 8: architecture of the lung (adapted from (Liu et al., 2006)). The major cell types composing the different lung epithelia are shown, together with representative stainings of the respective cell type-specific markers.

Cre recombinase expression under the control of promoters targeting Clara, NE and ATII cells (CC10, CGRP and SPC, respectively) in mice harboring conditional alleles of p53 and Rb1 demonstrated that NE cells are the cell lineage most likely to initiate SCLC, with a subset of ATII cells also giving rise to SCLC, and Clara cells being unable to initiate SCLC at all (Sutherland et al., 2011). Another report using a similar strategy showed that deletion of p53 and Rb1 in Clara or ATII cells does not lead to SCLC, and that SCLC probably arises only from NE cells (Park et al., 2011). These two studies are consistent with NE cells being the most frequent cell lineage at the origin of SCLC **(Figure 9)**, with apparent contradictory findings concerning the contribution of ATII cells, which could be attributed to the different strategies used to express the Cre recombinase (adenovirus *versus* transgenic mice), the fidelity of the promoters used to express Cre into a

particular cell lineage, the heterogeneity within the ATII cell population, or the genetic background of the mice (Blanpain, 2013).

As previously described, oncogenic mutations of the K-Ras gene are found in 25–50% of human lung adenocarcinomas. Introduction of constitutively active K-Ras^{G12D} in mice using intranasal infection led to lung adenocarcinomas, in which a small population of cells co-expressing Clara cell (CC10) and ATII cell (SPC) markers were identified (Jackson et al., 2001). In normal lungs, the so-called bronchioalveolar stem cells (BASCs), a population with stem cell properties located at the bronchioalveolar duct junctions (BADJs) also express these two markers, and expand upon K-Ras activation both *in vivo* and *in vitro*, arising the possibility that BASCs could work as cancer stem cells for lung adenocarcinoma (Kim et al., 2005) **(Figure 9)**. However, a recent study compared the oncogenic potential of KRas^{G12D} expression and p53 deletion in different airway cell populations using two different inducible Cre recombinases targeting ATII and the majority of BASCs (Sfpc-CreERT) or Clara cells and some ATII cells (CC10- CreERT). From these experiments it was found that BASCs do indeed proliferate in response to oncogenic K-Ras and form hyperplasias, but they do not progress into malignant adenocarcinomas (Xu et al., 2012). In contrast, K-Ras^{G12D} expression using Sfpc-CreERT in the ATII cells leads to adenocarcinoma formation, whereas no lesions were found to originate from the BASCs despite the high frequency of genetic recombination in these cells. Although the expression of oncogenic K-Ras from the two different promoters led to a histologically similar form of adenocarcinoma, some interesting differences (including different Sox2 expression) were observed between the ATII cells targeted by the CC10 and Sfpc-CreERT (Xu et al., 2012).

Overall, the studies conducted so far have started to shed light on the complex scenario of lung tumor origin, although the differences in the approaches used to artificially model human lung tumorigenesis in mouse makes it difficult to finally state whether a unique cell of origin for lung adenocarcinoma does actually exist, or whether different cell types can be induced to transform under different experimental conditions.

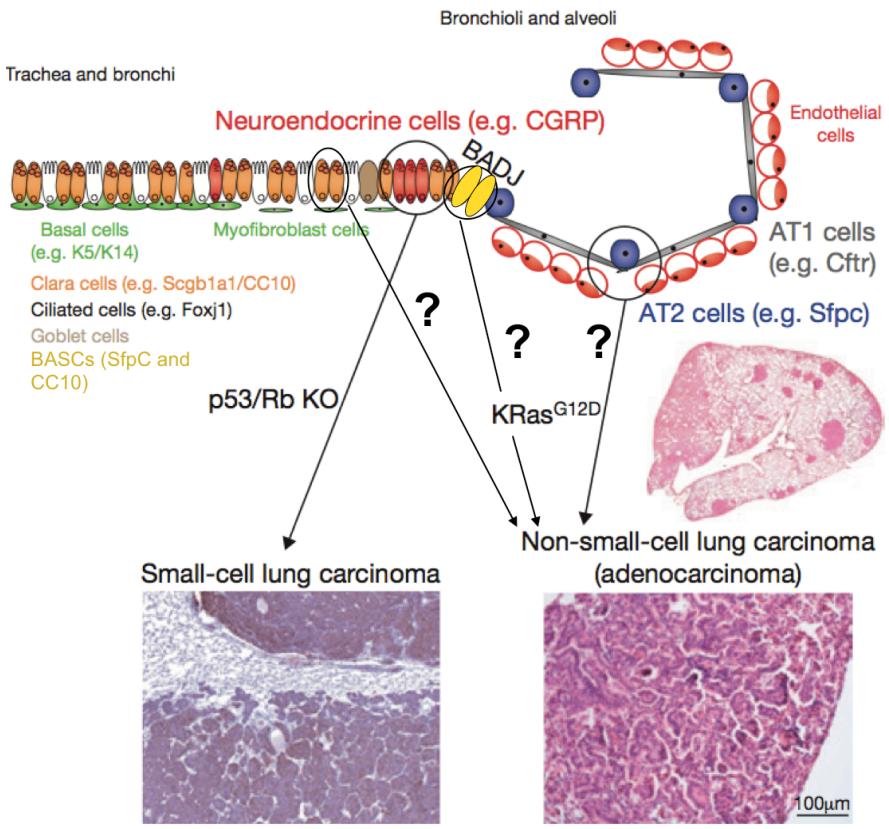


Figure 9: cells of origin in lung cancers (adapted from Blanpain, 2013). Schematic representation of different cell lineages in the lung. Neuroendocrine cells are the most sensitive cell lineage to initiate SCLC upon p53 and Rb deletion. The cell of origin of lung adenocarcinoma is more controversial.

3.4.2 CANCER STEM CELLS (CSCS) FOR LUNG CANCER?

While a cancer cell of origin is defined as the cell that upon accumulation of somatic mutations and other alterations undergoes transformation and starts uncontrolled proliferation, CSCs are considered as the only cell subpopulation, inside an already formed tumor, which is capable of indefinitely proliferate. In fact, the CSC model postulates that only a small subset of cells within a tumor is responsible for sustaining tumorigenesis and establishing the cellular heterogeneity characteristic of the primary lesion (Visvader and Lindeman, 2012); (Gupta et al., 2009) **(Figure 10)**. CSCs have been defined based on their ability to seed tumors in animal hosts, to self renew and to generate differentiated (non CSCs) progeny. Accordingly, the representation of CSCs within a tumor can be measured by the number of cells needed, at limiting dilutions, to seed new tumors (Clarke and Fuller, 2006). Pioneering work in this field originated from studies of leukemia stem cells and later included demonstrations of CSCs in solid tumors like human breast and brain cancers (Lapidot et al., 1994); (Al-Hajj et al., 2003); (Singh et al., 2003). On the contrary, the biology of lung cancer stem cells remains less well

studied. This difference is due in part to the lack of validated human lung stem cell markers as well as to the genotypic and histological diversity found in lung cancer (Sullivan et al., 2010). However it was recently discovered that the identity of the lung cancer stem cell population that promotes adenocarcinogenesis is likely dependent on the onco-genotype driving the malignancy (Curtis et al., 2010). In fact, the above mentioned BASCs, isolated from adenocarcinomas of different genotypes, were not always able to propagate the tumors when intratracheally transplanted into nude mice (Curtis et al., 2010), depending on the driving mutation of the tumor model. This raised the question whether these tumors are initiated and/or driven by another lung stem cell population (non-BASCs) or whether they are driven from a non-stem cell population. Moreover, these studies highlight the need for more reliable markers for lung tumor initiating/propagating cells

In summary, both the cell of origin and the tumor-sustaining cell population (if the CSC model can be applied) for lung adenocarcinoma still remain elusive, although some evidences suggest that different populations can serve as cancer initiating or cancer stem cells depending on the specific genetic alterations and probably other molecular characteristics of each peculiar tumor subtype. To date, a transcriptional profiling of early lung cancer cells, potentially able to provide these crucial informations, has not been reported and constitutes one of the main objectives of this work.

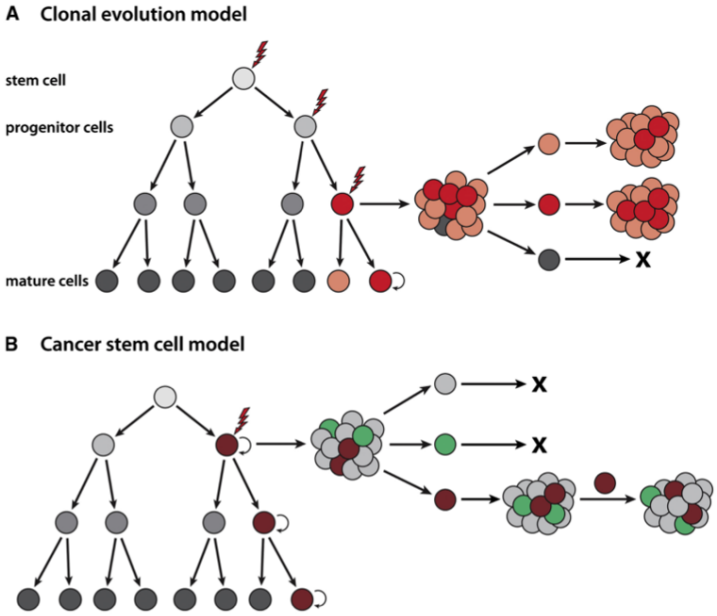


Figure 10: scheme of the clonal evolution and cancer stem cell models (adapted from Visvader and Lindman, 2012). **A**, clonal evolution: a cell (red) acquires mutations and produces a dominant clone. Some tumor cells (red and orange) arising from it have similar tumorigenic capacities while others (grey) may lack it due to stochastic events. **B**, a mutation in a progenitor cell (brown) endows it with stem cell-like properties. Only these cells have self-renewing capability and give rise to a range of different tumor cells (grey and green).

4 OBJECTIVES

THE OBJECTIVES OF THIS PHD THESIS WERE THE FOLLOWING:

1. To trace the expansion of lung cells following expression of the K-Ras oncogene in the K-Ras^{LSLG12Vgeo} mouse model of lung adenocarcinoma, in order to identify the cell type/s that give origin to the tumors.
2. To trace the expansion of lung cells following expression of the K-Ras oncogene during embryonic lung development, in order to study whether developmental timing can influence the susceptibility of different lung cell types to K-Ras^{G12V}-induced transformation.
3. To isolate and characterize at the gene expression level the cells that expand shortly after K-Ras^{G12V} activation .
4. To identify and validate possible markers of lung cancer initiating cells and eventually possible therapeutic targets.

ESTE TRABAJO PRESENTABA LOS SIGUIENTES OBJETIVOS:

1. Seguir la expansión de los varios tipos celulares en pulmón tras la activación del oncogen K-Ras en el modelo de adenocarcinoma K-Ras^{LSLG12Vgeo}, con el fin de identificar los tipos celulares capaces de originar tumores.
2. Seguir la expansión de los varios tipos celulares en pulmón tras la activación del oncogen K-Ras durante el desarrollo embrionario del pulmón, con el fin de estudiar como el estadio de desarrollo puede influenciar la predisposición de ciertos tipos celulares a la transformación inducida por K-Ras^{G12V}.
3. Aislar y caracterizar a nivel de expresión genica la población celular que se expande poco tiempo después de la activación del oncogen K-Ras.
4. Identificar y validar posibles marcadores de las celulas iniciadoras del cancer de pulmón, así como posibles dianas terapéuticas.

5 MATERIALS AND METHODS

5.1 Maintenance of mouse lines.

5.1.1 Proceeding of the mouse lines used in this work

The K-Ras^{LSLG12Vgeo}, RERT (Guerra et al., 2003) , the K-Ras^{lox} (Puyol et al., 2010) and the K-Ras^{FSFG12V} (Drosten, unpublished) lines were generated in Mariano Barbacid's laboratory (Spanish National Cancer Centre, Madrid, Spain). The Rosa26^{LSLLacZ} line (Mao et al., 1999) was generated in Stuart H. Orkins' laboratory (Dana Farber Cancer Institute, Boston, MA, USA). The Rosa26^{LSLEYFP} line (Srinivas et al., 2001) was generated in Frank Costantini's laboratory (Columbia University, New York, USA). The Tg^{Sca1Cre} and the Tg^{Sca1CreERT2} lines (Sánchez-García, unpublished) were generated in Isidro Sánchez-García's laboratory (CSIC/University of Salamanca, Salamanca, Spain). The Etv4^{NLacZ} line (Livet et al., 2002) was generated in Silvia Arber's laboratory (University of Basel, Basel, Switzerland).

5.1.2 Maintenance of mice

All animals used in this project were maintained in the animal facility of CNIO (Spanish National Cancer Centre) according to the FELASA (Federation of European Laboratory Animal Science Association) recommendations and following the European Union legislation. All the experiments described in this thesis were previously approved by the Bioethics and Animal Welfare Committee of the Institute for Health Care Carlos III. At CNIO animal facility mice are submitted to a light/darkness cycle of 12 hours each. The daytime cycle comes from fluorescent lamps that emit a white light (TLD 36W/840 and TLD 58W/840).

5.1.3 Necropsy

For the necropsy of adult animals:

1. The mouse was sacrificed in CO_2 chamber.
2. Various tissue samples were collected (pancreas, spleen, thymus, liver, kidney, lung, heart, skeletal muscle, intestine, stomach, colon, skin, ovary, uterus, prostate, testis, mammary gland, white adipose tissue, brown adipose tissue, brain, eye, femur were most commonly collected).

3. The samples were fixed in formalin.
4. Samples were processed, cut and stained with hematoxilin and eosin (H&E) by the Comparative Pathology Unit at CNIO.

Some organs were frozen in OCT (Sakura) and stored at -80 °C for future cryostat processing. Besides, some organs were also stained through X-Gal whole mount protocol (see section 5.3.1) or alkaline phosphatase whole mount protocol (see section 5.3.5). In this case they were fixed in the adequate solution.

Embryonic tissues were collected as follows at different time points during gestation:

1. The mother was sacrificed in CO_2 chamber.
2. The embryos were extracted from the uterus.
3. The embryos were dissected and lungs were taken (from E9.5 to E13.5 the dissection was done under a magnifying optical instrument).
4. The embryonic lungs were fixed in X-Gal fixation solution (see section 5.3.1) for staining.

5.1.4 Genotyping

The genotyping of mice was routinely made by extracting DNA from the mice tails.

1. 490 µl of lysis buffer (20mM Tris/HCl pH8.0, 100mM NaCl, 0.5% SDS, 10mM EDTA pH8.0 and milli-q H_2O) and 10 µl of proteinase K were added to each tail, and left o.n. at 55°C.
2. 300 µl of saturated NaCl were added.
3. The solution was mixed and incubated for 10 min at 4°C.
4. The samples were centrifuged at 13000 rpm for 10 min.
5. The supernatant was transferred to a clean eppendorf.
6. 800 µl of isopropanol were added, then the solution was mixed vigorously.
7. The mixture was incubated for at least 1 hour at -20°C.
8. The samples were centrifuged at 13000 rpm for 10 min.
9. The supernatant was discarded.
10. 400 µl of 70% ethanol (EtOH) were added.
11. The samples were centrifuged at 13000 rpm for 10 min.

12. The supernatant was discarded.
13. The pellets were left to dry.
14. The pellets were resuspended in 100 µl of milli-q H_2O.

Once the DNA was isolated we performed the genotyping of the mice using polymerase chain reaction (PCR) technology. Adding for each reaction:

- 1.25 µl $MgCl_2$ 25mM.
- Taq Polymerase Buffer 1X.
- 0.5 µl dNTPs 10mM.
- 0.1 µl BSA 10 mg/ml.
- 0.1 µl Taq Polymerase (5 u/µl EcoTaq, Ecogen).
- 1 µl of each of the primers (10µM, Sigma).
- 1 µl of DNA.
- Fill up to 20 µl with milli-q H_2O.

The oligonucleotides used for genotyping were:

K-Ras^{LSLG12Vgeo}

K-Ras2F_16B5: 5' CGTCCAGCGTGTCCTAGACTTTA 3'

K-Ras2R_15B9: 5' ACTATTTCATACTGGGTCTGCCTT 3'

NeoF_2B1: 5' TGACCGCTTCCTCGTGCTT 3'

KI allele: K-Ras2R_15B9-NeoF_2B1, 390bp.

WT allele: K-Ras2F_16B5-K-Ras2R_15B9, 240bp.

K-Ras^{lox}

K-Ras2F_8B2: 5' CCACAGGGTATAGCGTACTATGCAG 3'

3'Ex1: 5' CTCAGTCATTTTCAGCAGGC 3'

KI allele: 350bp.

WT allele: 550bp.

K-Ras^{FSFG12V}

3'Ex1: 5' CTCAGTCATTTTCAGCAGGC 3'

STOP: 5' TAGTGCCTTGACTAGAGATCA 3'

K-Ras2F_8B2: 5' CCACAGGGTATAGCGTACTATGCAG 3'

KI allele: STOP- 3'Ex1, 507bp.

WT allele: 3'Ex1- K-Ras2F_8B2, 350bp.

RERT

Polr2aR_10B6: 5' CCTCTCTGAGCCTCAATTAAGCAG 3'

ESRF1f_10B7: 5' TGAGTAACAAAGGCATGGAGCA 3'

Polr2aF_14B5: 5'CCAGATGACAGCGATGAGGA 3'

KI allele: Polr2aR_10B6-ESRF1f_10B7, 390bp.

WT allele: Polr2aR_10B6-Polr2aF_14B5, 480bp.

Rosa26^{LSLEYFP}

oIMR0316: 5' GGAGCGGGAGAAATGGATATG 3'

oIMR0883: 5' AAAGTCGCTCTGAGTTGTTAT 3'

oIMR4982: 5' AAGACCGCGAAGAGTTTGTC 3'

KI allele: oIMR0316-oIMR4982 ,320bp.

WT allele: oIMR0316-oIMR0883, 600bp.

Rosa26^{LSLLacZ}

oIMR8052: 5' GCGAAGAGTTTGTCCTCAACC 3'

oIMR8545: 5' AAAGTCGCTCTGAGTTGTTAT 3'

oIMR8546: 5' GGAGCGGGAGAAATGGATATG 3'

KI allele: oIMR8052-oIMR8545, 340bp.

WT allele: oIMR8545-oIMR8546, 650bp.

Etv4^{NLacZ}

454: 5’ CAGCCTCTGTTCCACATACACTCC 3’

o91: 5’ TAGTATCGCAGCGAGCGGCTCAGC 3’

75: 5’ GGAATCTTGGGCCTTGAGAACAGC 3’

77: 5’ GTGTGATGTACATATGCCCTAACC 3’

KI allele: 454-o91, 450bp.

WT allele: 75-77, 686bp.

Tg^{Sca1Cre} and **Tg^{Sca1CreERT2}**

TETO1: 5’ CCGGTTATTCAACTTGCACC 3’

TETO2: 5’-CTGCATTACCGGTCGATGCAAC-3’

149bp.

The PCR program used was:

94°C 1 min

94°C 30 sec
60°C 30 sec } x 35 cycles
72°C 1 min

72°C 10 min

The length of the amplified fragments was assessed by 2% agarose gel electrophoresis. Alternatively, the mice were genotyped by Transnetyx (Córdova, TN, USA).

5.2 *In vivo* proceedings

5.2.1 Tamoxifen (4-OHT) treatment

The 4-OHT has a steroid type chemical structure that makes it only soluble in organic solvents, which cannot be used as a vehicle for intraperitoneal administration due to

significant toxic effects. For administration of 4-OHT, the steroid was dissolved in corn oil.

1. 50 mg of 4-OHT (Sigma) were dissolved in 10 ml of corn oil (Sigma).
2. The mixture was sonicated until the 4-OHT was completely dissolved (2 min at 42% amplitude).
3. 200 µl of the dissolved 4-OHT were administered through intraperitoneal injection to each mouse.

The duration of the treatment varied depending on the mouse strain and on the type of experiments. For tracing of the K-Ras^{G12V}-expressing cells, a single injection was given at P21. To induce lung tumors in Tg^{Sca1CreERT2}; K-Ras^{LSLG12Vgeo} line, mice were subjected to a 2 weeks treatment, injecting 3 times per week.

5.2.2 Adenoviral intratracheal infection

All the adenoviral preparations used in this work were purchased from the Iowa University (Iowa City, Iowa, USA).

Mice were treated when 8 weeks old with a single inoculation of 30 µl of the adenoviral solution (2.5x10^8 pfu/mouse for adeno-Cre, 6x10^7 pfu/mouse for adeno-Flipase). Mice were previously anhestesized via intraperitoneal injection of a ketamine (75 mg/kg, Imalgene®) xilacine (12 mg/kg, Rompum®) solution. The virus was introduced directly in the trachea through a cannula (Simpson et al., 2001). After inoculation mice were maintained during 10-15 minutes under a heat lamp.

5.3 Processing of mouse tissues

5.3.1 X-Gal whole mount staining

1. Lungs and other organs were obtained during the necropsy of the mice.
2. The tissues were fixed at RT for 60-90 min in:
   - 0.2% glutaraldehyde (Sigma)
   - 1.5% formalin solution (Sigma)

- 2mM $MgCl_2$
- 5mM EGTA
- 100mM sodium phosphate pH7.3
- milli-q H_2O

3. The tissues were washed 3 times (20 min each at RT) in washing solution:
   - 0.2% NP-40
   - 0.1% sodium deoxycolate (Sigma)
   - 2mM $MgCl_2$
   - 100mM sodium phosphate pH7.3
   - milli-q H_2O
4. Following the washes, the tissues were stained for 48 hours at 37°C in the staining solution:
   - 0.2% NP-40
   - 0.1% sodium deoxycolate (Sigma)
   - 2mM $MgCl_2$
   - 100mM sodium phosphate pH7.3
   - 5mM $K_3Fe(CN)_6$ (ProLab)
   - 5mM $K_4Fe(CN)_6$ (ProLab)
   - 1mg/ml X-Gal (dissolved in dimethylformamide)
   - milli-q H_2O
5. The tissues were washed 3 times and for 10 min with the above mentioned washing solution.
6. The samples were postfixed o.n. in phosphate buffered 10% formaldehyde.
7. The fixed and stained tissues were washed twice with 1X PBS for 5 min.
8. The tissues were washed twice with 50% EtOH and twice with 70% EtOH (1 hour each).
9. At this point tissues were processed by the Comparative Pathology Unit at CNIO, which produced the standard 4 µm sections on a 3-aminopropyltrethoxylan coated slides and counter-stained with Nuclear Fast Red (NFR).

5.3.2 X-Gal staining on criostate sections

1. The frozen OCT samples were taken out from -80°C freezer.
2. The samples were cut in the cryostat in 10 µm thick sections.
3. The cut samples were placed on a slide.
4. The slides were stored at 4°C o.n. or 1 hour at RT.
5. Sections were washed with PBS 1X.
6. Sections were fixed for 10 min at RT in the following solution:
   - 0.2% glutaraldehyde (Sigma)
   - 5mM EGTA
   - 2mM $MgCl_2$
   - 0.1M phosphate buffer pH7.3
   - milli-q H_2O
7. The sections were washed 3 times with the following washing solution:
   - 2mM $MgCl_2$
   - 0.1M phosphate buffer pH7.3
   - 0.02% NP-40
   - 0.01% sodium deoxycolate (Sigma)
   - milli-q H_2O
8. After the washes, the sections were stained o.n. at 37°C with:
   - 2mM $MgCl_2$
   - 0.1M phosphate buffer pH7.3
   - 0.02% NP-40
   - 0.01% sodium deoxycolate (Sigma)
   - 1mg/ml X-Gal
   - 5mM $K_3Fe(CN)_6$ (ProLab)
   - 5mM $K_4Fe(CN)_6$ (ProLab)
   - milli-q H_2O
9. The sections were washed 3 times with the washing solution previously described.

10. The counter-staining was performed with NFR.

5.3.3 Nuclear Fast Red

1. The NFR solution was prepared:
   - 25 gr aluminium sulphate (Sigma)
   - 0.5 gr NFR (Sigma)
   - 500 ml milli-q H_2O
   - boil the solution for 2-3 min
   - filter the solution
2. The samples were washed with H_2O.
3. The sections were stained with NFR for 1:30 min.
4. The sections were washed with H_2O.
5. The sections were washed twice for 2 min with 70% EtOH.
6. The sections were washed twice for 2 min with 100% EtOH.
7. The sections were rinsed with Xilol (Merck) for 1 min.
8. The sections were mounted with the coverslip.

5.3.4 X-Gal destaining

1. The deparafinized sections from whole mount X-Gal stained lungs were washed twice with 70% EtOH during 10 min.
2. The sections were washed twice with 100% EtOH during 10 min.
3. The sections were incubated with Xilol until the blue color was dissolved.
4. The sections were washed twice with 100% EtOH during 10 min.
5. The sections were washed twice with 70% EtOH during 10 min.
6. The sections were put in PBS.

5.3.5 Immunofluorescence

1. The sections were deparaffinized.
2. The deparaffinized sections were treated with sodium citrate 10mM (pH6) at 95°C for 20 min (antigen retrival).

3. The samples were permeabilized for 30 min at RT with:
- 0.4% Triton
- 0.04% SDS
- 1X PBS
4. The slides were washed twice with 1X PBS.
5. The slides were blocked for 45 min in 5% BSA at 37°C.
6. The primary antibodies were diluted in 5% BSA (rabbit anti-SP-C 1:50 Millipore, goat anti-SP-C 1:50 Santa Cruz, goat anti-CC10 1:50 Santa Cruz, rabbit anti-AQPV 1:50 Abcam, rabbit anti-CGRP 1:50 Sigma) and incubated with the samples 1 hour at 37°C.
7. The slides were washed twice with 1X PBS.
8. The secondary antibodies were diluted in 5% BSA (Alexa Fluor 555 donkey anti-goat 1:200, Alexa Fluor 647 donkey anti-rabbit 1:200, Invitrogen) and incubated with the samples 45 min at 37°C.
9. The first wash was done with diluted DAPI 1:1000 in 1X PBS.
10. Two additional washes were performed with 1X PBS.
11. The slides were mounted with mowiol and the coverslip.
12. The slides were ready to be analyzed at the confocal microscope Leica SP2.

5.3.6 Histopathology and immunohistochemistry

Tissues including pancreas, spleen, thymus, liver, kidney, lung, heart, skeletal muscle, intestine, stomach, colon, skin, ovary, uterus, prostate, testis, mammary gland, white adipose tissue, brown adipose tissue, brain, eye, femur were dissected, fixed in 10%-buffered formalin (Sigma) and embedded in paraffin. 4 µm tick sections were cut and stained with hematoxylin and eosin (H&E). Antibodies used for immunohistochemistry included rabbit polyclonal anti-SP-C (Millipore), goat polyclonal anti-CC10 (Santa Cruz), rabbit polyclonal anti-CGRP (Sigma), rabbit polyclonal anti-Etv4 (Epitomics), rat monoclonal anti-β-galactosidase (produced by the Monoclonal Antibody Unit at CNIO), mouse monoclonal anti-bromo-deoxyuridine (GE Healtcare), rabbit polyclonal anti-caspase 3 active (R&D Systems), rabbit polyclonal anti-cytokeratin 5 (Covance), rabbit

monoclonal anti-Ki67 (Master Diagnostica), mouse monoclonal anti-p63 (Thermo Scientific). For detection of β-galactosidase activity in adult and embryonic tissues, samples were included in OCT compound (Sakura) and frozen. X-Gal staining of 10 µm tick criosections was performed as previously described (section 5.3.3).

5.4 Quantification of tumors and pretumoral lesions

5.4.1 Quantification of the relative areas occupied by the different types of lung epithelia

In order to estimate the realtive areas occupied in a normal lung by the different types of epithelia (alveoli, bronchioles and bronchioalveolar duct junctions – BADJs), we scanned H&E stained 2D sections from paraffin-embedded wild type lungs (n=3, one section per mouse) and we calculated the areas based on morphologic features using the Panoramic Viewer software (3DHISTECH).

5.4.2 Quantification of K-Ras^{G12V}-induced hyperproliferative clusters at early time points.

Quantification of the blue β-Gal-positive area at different time points during K-*Ras*^{G12V}-induced proliferation was performed using the Panoramic Viewer software (3DHISTECH) after scanning random 2D sections from X-Gal whole mount stained lungs of K-Ras$^{lox/LSLG12Vgeo}$ RERT$^{ert/ert}$ mice.

Quantification of K-Ras^{G12V}-positive single cells and small clusters at very early time points (1, 2 or 4 weeks) was performed manually in random sections from X-Gal whole mount stained lungs of K-Ras$^{lox/LSLG12Vgeo}$ RERT$^{ert/ert}$ mice.

In order to quantify the distribution of the K-Ras^{G12V}-positive clusters in the different lung areas (alveoli, bronchioles and BADJs) according to the cluster size (4-20 cells or >20 cells) we scanned 2D sections from X-Gal whole mount stained lungs of 6 K-Ras$^{lox/LSLG12Vgeo}$ RERT$^{ert/ert}$ mice who recieved a single 4OHT injection and were sacrificed

after 1month. The estimation of the cluster size and annotation of the localization were done manually.

For the quantification of the different lung cell types in the K-Ras^{G12V}-positive clusters at early time points, the blue clusters were localized in sections from X-Gal whole mount stained lungs from 2 mice sacrificed 4 weeks after 4OHT injection (a total of 45 β-Gal-positive clusters were quantified). The consecutive section was then destained from the X-Gal precipitates by xilol incubation (see section 5.3.4) and subsequently stained by immunofluorescence (IF) with antibodies against SPC and CC10. β-Gal-positive groups were clustered by size (4-20, 20-60, >60 cells) and the cell type expressing the LacZ reporter was assessed in the consecutive section according to the IF staining.

5.4.3 Tumor quantification

Quantification of K-Ras^{G12V}-induced lesions at different time points was performed analyzing serial sections from either X-Gal whole mount stained lungs or formalin-fixed unstained lungs. Sections number 1, 10, 20, 30, 40, 50, 60, 70, 80, 90 and 100 were counterstained with either NFR or H&E and the number of adenomas and adenocarcinomas was assessed manually based on histopathologic criteria.

Quantification of SPC and CC10 positive lesions in both 4OHT-treated or adeno-Cre infected mice 6 months after K-Ras^{G12V} induction was performed analyzing serial sections from paraffin-embedded lungs. Sections number 9, 19, 29, 39, 49, 59, 69, 79, 89 and 99 were stained with anti-SPC antibody by immunohistochemistry, while sections number 10, 20, 30, 40, 50, 60, 70, 80, 90 and 100 were stained with anti-CC10 antibody.

5.5 FACS sorting

5.5.1 Lung single cells preparation

1. Mice were anesthetized with an intraperitoneal injection of 6.5µl/gr of the following solution:
    - 700 µl physiological serum
    - 300 µl medetomidine (Domtor®)

- 1.5 ml ketamine (Imalgene®)

2. The abdominal cavity was opened and mice were bled by severing the inferior vena cava.
3. Mice were perfused injecting 20 ml of physiological serum through the left ventricle of the heart.
4. 2 ml dispase (50 units/ml, BD Biosciences) were injected through the trachea.
5. The lungs were removed and and put in a falcon tube with 5 ml HBBS + antibiotics + 1%BSA.
6. Under a flow hood in the cell culture room, the lungs were transferred in a Petri dish and minced with a scalpel.
7. The minced lungs were transferred into a GentleMacs (Milltenyi Biotec) tube with 7 ml of HBBS + antibiotics + 1%BSA + 60units/ml DNase1 + 70units/ml Collagenase type 1.
8. The tubes were put in the GentleMacs machine with the program “mouse heart 1”.
9. The tubes were incubated 30 min at 37°C rotating.
10. The tubes were put in the GenletMacs machine with the program “mouse heart 2”.
11. The tubes were incubated 30 min at 37°C rotating.
12. The tubes were put again in the GentleMacs machine with the program “mouse heart 2”.
13. The cell preparations were filtered through a 100 µm stainer, then through a 40 µm one.
14. The cells were centrifuged 5 min at 1200 rpm.
15. The supernatant was discarded and the cells resuspended in 2 ml ACK Lysis Buffer (Lonza) to lyse the red blood cells. The cells were incubated 4 min at RT, then serum-free DMEM medium was added to stop the reaction and cells were centrifuged.

5.5.2 Staining of live lung cells

1. The cells were resuspended in 10 ml 3% BSA and incubated 30 min on ice for blocking.
2. The cells were centrifuged and resuspended in 1ml of the antibody solution (depending on the experiments the antibodies used were: anti-CD45-APC, anti-CD45-FITC, anti-CD31-APC, anti-CD31-FITC, BD Pharmingen, 1:800; PerCp-Cy5.5 anti-Sca1, BD Pharmingen, 1:400). The cells were incubated during 15 min on ice.
3. PBS was added up to 13 ml and the cells were centrifuged.
4. The pellets were resuspended in sorting buffer (1X PBS + 1mM EDTA + 25mM Hepes + 3%FBS) + DAPI 1:1000.

5.5.3 Fluorescence Activated Cell Sorting

The stained cells were analyzed using a FACSAria II or a INLFUX machine from BD Biosciences. Debris and doublets were first excluded based on the forward and side scatter parameters. Dead cells were then excluded by means of DAPI incorporation. The results were analyzed using the FlowJo software. For subsequent RNA extraction, cells were directly sorted into BLT buffer supplied with β-mercaptoethanol (see section 5.6).

5.6 Quantitative Real Time RT-PCR

5.6.1 RNA extraction

Total RNA was extracted from sorted samples (see section 5.5) and tissues using RNeasy Micro Kit (Quiagen). Briefly, the cells were lysed by suspending them into BLT buffer supplied with β-mercaptoethanol and by passing them 5 times through a 20-gauge needle. To disgregate tissues, a homogenizer was used. The RNA was then purified using the appropriate columns following the manufacturer's instructions. On-column DNAase digestion was performed to eliminate residual genomic DNA. RNA was eluted in 20 µl of elution buffer. Finally, RNA was quantified using Nano-drop spectrophotometer. Also, to assess the quality of RNA for subsequent gene expression microarray experiments, the

Agilent 2100 Bioanalyzer was used: only RNAs with a RNA integrity number (RIN) ≥7 were chosen for subsequent microarray hybridization.

5.6.2 cDNA synthesis

cDNA synthesis was performed using SuperScript II RT kit from Invitrogen.

1. For each sample the following mix was prepared:
    - 200 ng random primers
    - 40-100 ng total RNA
    - 1 µl dNTP mix (10mM each)
    - Sterile, distilled H_2O
2. The mixture was incubated at 65°C for 5 min.
3. The samples were chilled on ice and spinned.
4. For each sample the following mix was added:
    - 4 µl 5X First Strand Buffer
    - 2 µl 0.1M DTT
    - 1 µl RNaseOUT (40 units/µl)
5. The content of the tubes was gently mixed. Samples were incubated at 25°C for 2 min.
6. 1 µl of SuperScript II RT enzyme was added and samples were mixed by pipetting.
7. Samples were incubated at 25°C for 10 min.
8. Samples were incubated at 42°C for 50 min.
9. The reaction was inactivated by heating at 70°C for 10 min

5.6.3 Real Time PCR

5.6.3.1 Primer design criteria

1. Tm: 58°C to 61°C.
2. Primer length: 19-24 bp
3. Guanine cytosine (GC) content: 45-55%
4. PCR amplicon length: 100-200 bp

5. Cover an exon-exon junction (to avoid the problem of possible residual genomic DNA contamination).
6. Order primers with standard desalting purification.

5.6.3.2 Primers

Etv4

Etv4 forward: 5'CGACTCAGATGTCCCTGGAT 3'

Etv4 reverse: 5'GCCTGTCCAAGCAATGAAAT 3'

Etv1

Etv1 forward: 5'AGCAGAACAGAAGGCTGCAT 3'

Etv1 reverse: 5'CTTGCTTAATGTCCCCCATCG 3'

Etv5

Etv5 forward: 5'GCCATGAAGGATTCCCGTAT 3'

Etv5 reverse: 5'CTTTGCCTTCCAGTCTCTCG 3'

GAPDH

GAPDH forward: 5'TGCACCACCAACTGCTTAG 3'

GAPDH reverse: 5'GGATGCAGGGATGATGTTT 3'

β-actin

β-actin forward: 5'GACGGCCAGGTCATCACTATT 3'

β-actin reverse: 5'AGGAAGGCTGGAAAAGAGCC 3'

5.6.3.3 qRT-PCR

1. The PCR mix was prepared as follows:
    - 10 µl SYBR Green Mix 2X (Applied Biosystems)
    - 0.5 µl primer forward 10µM
    - 0.5 µl primer reverse 10µM
    - 1 µl template cDNA
    - 8 µl H_2O
2. The samples were prepared in a 96-well plate.
3. The plate was sealed with optical adhesive film and spinned in the centrifuge.

4. The plate was introduced in the AB 7900HT machine.
5. The thermal cycle selected was the following:
    - Stage 1: 2 min at 50°C
    - Stage 2: 10 min at 95°C
    - Stage 3: 15 sec at 95°C then 1 min at 60°C (repeat 40 cycles)
    - Dissociation stage: 95°C for 15 seconds, 60°C for 15 seconds, followed by a slow ramp to 95°C

5.6.3.4 Data analysis

The data were analyzed using the SDS software (Applied Biosystems). The relative mRNA expression was calculated using the comparative Ct (cycle threshold) method:

[delta][delta]Ct = [delta]Ct,sample - [delta]Ct,reference

[delta]CT,sample is the Ct value for any sample normalized to the endogenous housekeeping gene and [delta]Ct, reference is the Ct value for the calibrator also normalized to the endogenous housekeeping gene.

5.7 Gene Expression Microarray

5.7.1 Microarray Hybridization

Single-color gene expression profiles were generated using the 60K format Whole Mouse Gene Expression Microarray from Agilent. Labeling and hybridization was performed following the manufacturer's protocol. In brief, 100 ng of RNA was linearly amplified and labeled with Cy3 using Agilent's one-color Quick Amp Labeling Kit following the instructions of the protocol. Then, 1650 ng of Cy3-labeled cRNA was hybridized on the 60K arrays using Agilent's High-RPM Gene Expression Hyb Kit. Hybridization was performed for 17h at 65°C in a rotating hybridization oven at 10 rpm according to the company's recommendations. After washing and scanning, resulting TIFF-images were processed using Agilent's Feature Extraction software Version 9.

5.7.2 Microarray analysis

Significance Analysis of Microarray was performed to generate differential expression data. Twofold and greater changes in gene expression were considered significant. Treewiew software was used to generate the heatmap. Gene Set Enrichment Analysis (GSEA) was performed to analyze pathways included in the major databases (Biocarta, GenMAPP, KEGG, Reactome).

5.7 Western Blot

5.7.1 Protein extraction

1. Either tissues or cells were resuspended in lysis buffer: 50mM Tris-HCl (pH7.4) solution, containing 150mM NaCl, NP-40 0.5% and protease (PMSF 100mM, Roche) and phosphatase inhibitors (100mM vanadate sodium, Roche and 1mM NaF, Sigma).
2. For tissue protein extraction we used a homogenizer to mechanically breakdown the tissue samples and maximize the efficiency of lysis.
3. The protein extracts were sonicated for 5 to 10 seconds at 10% amplitude.
4. The samples were incubated on ice for 15 min.
5. The extracts were centrifuged for 15 min at 13000 rpm to remove the undigested membranes.
6. The supernatant was transferred to a new tube.
7. In order to quantify the amount of protein obtained, we used the Bradford method (1 µl of the protein extract was added to 1 ml of Bradford reagent and the absorbance was read on a spectrophotometer).

5.7.2 Gel Run

1. The quantity of tissue/cell extracts needed to load the same total amount of protein was estimated (30-50 µg).
2. 4X loading buffer and 20X reducing agent (Biorad) were added.
3. The samples were boiled for 5 min. at 95°C.

4. The samples were placed on ice for 5 min.
5. Criterion XT precast gels (Biorad) were prepared into the cassette with running buffer (MES 1X, Biorad).
6. The samples were spinned and loaded into the gel.
7. 10 µl of the molecular weight color marker were loaded as well.
8. The gel was run with a constant voltage of 120V for usually 1.5 hours.

5.7.3 Transfering the denatured proteins to a membrane

1. A piece of nitrocellulose transfer membrane was cut as well as six pieces of Whatman paper.
2. The membrane and the papers were dipped on transfer buffer.
3. The gel was also dipped in transfer buffer.
4. The transfer "sandwich" was assembled in a wet blotter.
   -3 Whatman papers- membrane- gel- 3 Whatman papers.
5. The gel was transferred for 30 min at constant 0.35 A.
6. When the transfer was finished, the amount of protein transferred was checked by staining with Ponceau S solution (Sigma).

5.7.4 Blocking and antibodies

1. The membrane was incubated in the blocking solution (5% BSA in 1X TBS-T) for 1 hour at RT.
2. The membrane was incubated with the primary antibody (rabbit anti-Etv4, Epitomics, 1:250; mouse anti-βactin, Sigma, 1:5000) diluted in the blocking solution o.n. at 4°C.
3. The membrane was washed twice with 1X TBS-T.
4. The membrane was incubated with the appropriate HRP-conjugated secondary antibody (1:2000, Dako).
5. The membrane was washed twice with 1X TBS-T.
6. Protein visualization was carried out with ECL detection reagent (Amersham) using different exposure times depending on the antibody.

5.8 Statystics

The statistics calculations reported in this work were performed using the 2-paired Student's t-test. Equality or difference in the variance was determined by the Fischer's F test. All the calculations were performed with the Excel software (Microsoft).

6 RESULTS

6.1 Lung cells are uniquely permissive to transformation by K-Ras oncogenes in adult mice.

Unbiased expression of the K-Ras^{G12V} oncogene in mouse adult tissues does not result in wide spread induction of proliferation or malignant transformation (Guerra et al., 2003). A notable exception to these findings is lung tissue. Indeed, adult K-Ras$^{+/LSLG12Vgeo}$; RERT$^{ert/ert}$ mice exposed to a tamoxifen diet for three months only develop overt NSCLCs in spite of expressing the K-Ras^{G12V} oncoprotein in more than 50% of their cells, as determined by expression of β-Gal, its surrogate marker. Some tissues such as the gastric epithelium also developed premalignant lesions including papillomas, but none of them displayed overt tumors even at one year of age.

These observations indicate that lung cells are uniquely permissive for K-Ras oncogene mediated transformation. In an effort to understand the nature of these cells and to shed light on the earliest steps of lung tumor development, we submitted K-Ras$^{lox/LSLG12Vgeo}$; RERT$^{ert/ert}$ mice to a single dose of 4OHT in order to activate the concomitant expression of the K-Ras^{G12V} oncoprotein and its surrogate marker, the bacterial β-Gal protein that can be easily detected by X-Gal staining. 4OHT was used at limiting dilution in order to activate K-Ras^{G12V} oncogene expression in a narrow number of cells that would allow us to follow their individual fate during the first rounds of cell division, before they yield anatomically identifiable lesions such as hyperplasia or adenomas. The model we used mimics the loss of heterozigosity (LOH) that frequently occurs in human K-Ras mutant lung tumors ((Li et al., 2003). For a detailed description of the model see section 3.3.1). While giving the same type of lesions as non-LOH K-Ras^{G12V}-mutant mice, the K-Ras$^{lox/LSLG12Vgeo}$ model allows faster tumor development and progression (Puyol et al., 2010).

Under the experimental conditions selected for this study, 0.1 to 0.2% of all lung cells became positive for X-Gal staining one week after 4OHT treatment. Importantly, these cells were uniformly distributed throughout the four main structures of the lung, bronchi, bronchioles, bronchioalveolar duct junctions (BADJ) and alveoli, indicating that Cre-mediated recombination leading to expression of the resident K-Ras^{G12V} oncogene

occurred in an unbiased, uniform fashion. These X-Gal positive cells rapidly acquired proliferative properties covering almost 8% of the total lung area three months after 4OHT exposure and 15% by six months **(Figure 11)**. Moreover, treated mice developed an average of 5 adenomas at three months (n=5) and 20 at six months (n=3) after 4OHT injection **(Figure 12)**. Likewise, they displayed few adenocarcinomas five to six months after receiving 4OHT **(Figure 12)**.

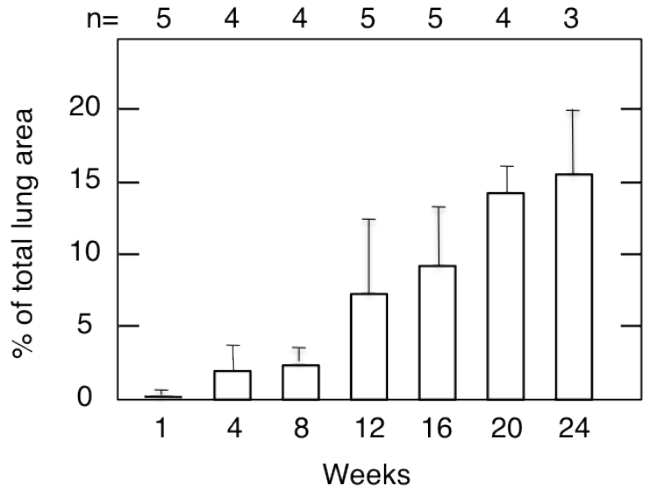


Figure 11: Expansion of the β-Gal-positive lung area at different time points after K-Ras^{G12V} activation. K-Ras$^{lox/LSLG12Vgeo}$; RERT$^{ert/ert}$ mice received a single dose of 4OHT and were sacrificed at the indicated time points (on the *x* axes) after treatment. n= number of mice.

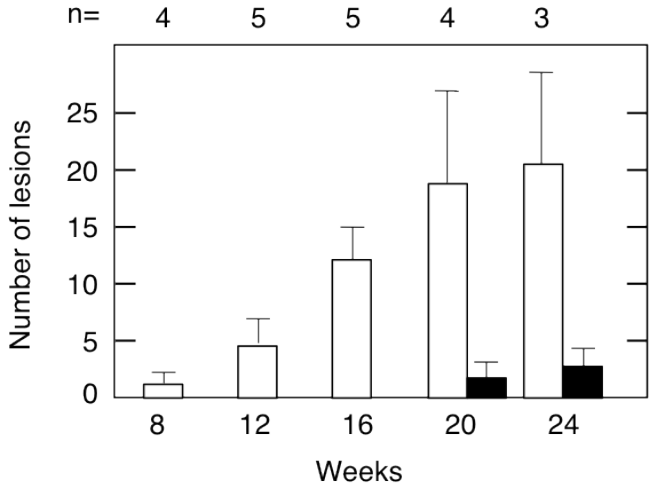


Figure 12: Increase in number and aggressiveness of lesions at different time points after K-Ras^{G12V} activation. White bars, adenomas; black bars, adenocarcinomas. K-Ras$^{lox/LSLG12Vgeo}$; RERT$^{ert/ert}$ mice received a single dose of 4OHT and were sacrificed at the indicated time points (on the *x* axes) after treatment. n= number of mice.

6.2 Adult alveolar type II cells are the predominant lung cancer initiating cells following unbiased K-Ras oncogene activation.

When mice were sacrificed one week after 4OHT injection, most K-Ras^{G12V} expressing X-Gal positive cells (>98%) appeared as isolated single cells indicating that they had not undergone cell division **(Figure 13)**. However, when mice were sacrificed two weeks after exposure to 4OHT, about 30% of the X-Gal positive cells appeared in clusters of 4 to 8 cells (groups of two cells were not counted due to the high variability) **(Figure 13)**. No significant differences were observed in the distribution among the different areas of the lung (**Figure 14** upper panel). These findings indicate that a large proportion of adult lung cells, regardless of their developmental lineage, have the capacity to initiate a proliferative response upon expression of an endogenous K-Ras^{G12V} oncoprotein. Thus, it is unlikely that these adult cells correspond to either stem or precursor cells. Comparison with Rosa26$^{LSLLacZ/LSLLacZ}$; RERT$^{ert/ert}$ reporter mice treated with the same dose of 4OHT confirmed that expansion of the K-Ras^{G12V}-positive cells was indeed due to the oncogene expression, as the cells targeted by the Cre recombinase in a K-Ras WT background failed to expand **(Figure 15)**.

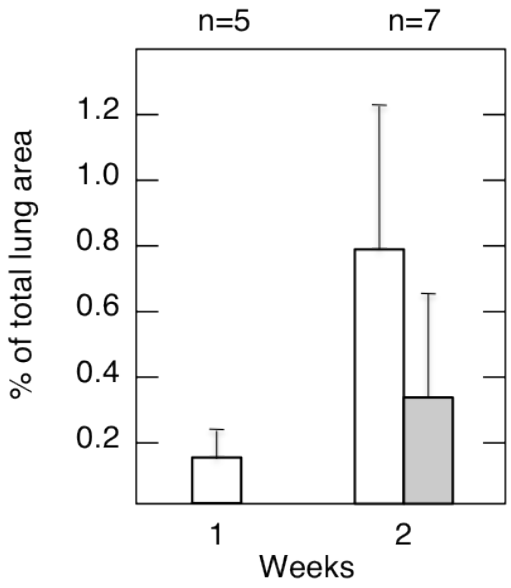


Figure 13: **Only a fraction of K-Ras^{G12V}-expressing cells in the lung undergo oncogene-induced proliferation**. Quantification of K-Ras^{G12V}-positive single cells (white bar) and clusters > 4 cells (grey bar) in K-Ras$^{lox/LSLG12Vgeo}$; RERT$^{ert/ert}$ mice treated with a single 4OHT injection and sacrificed either 1 or 2 weeks later. n= number of mice.

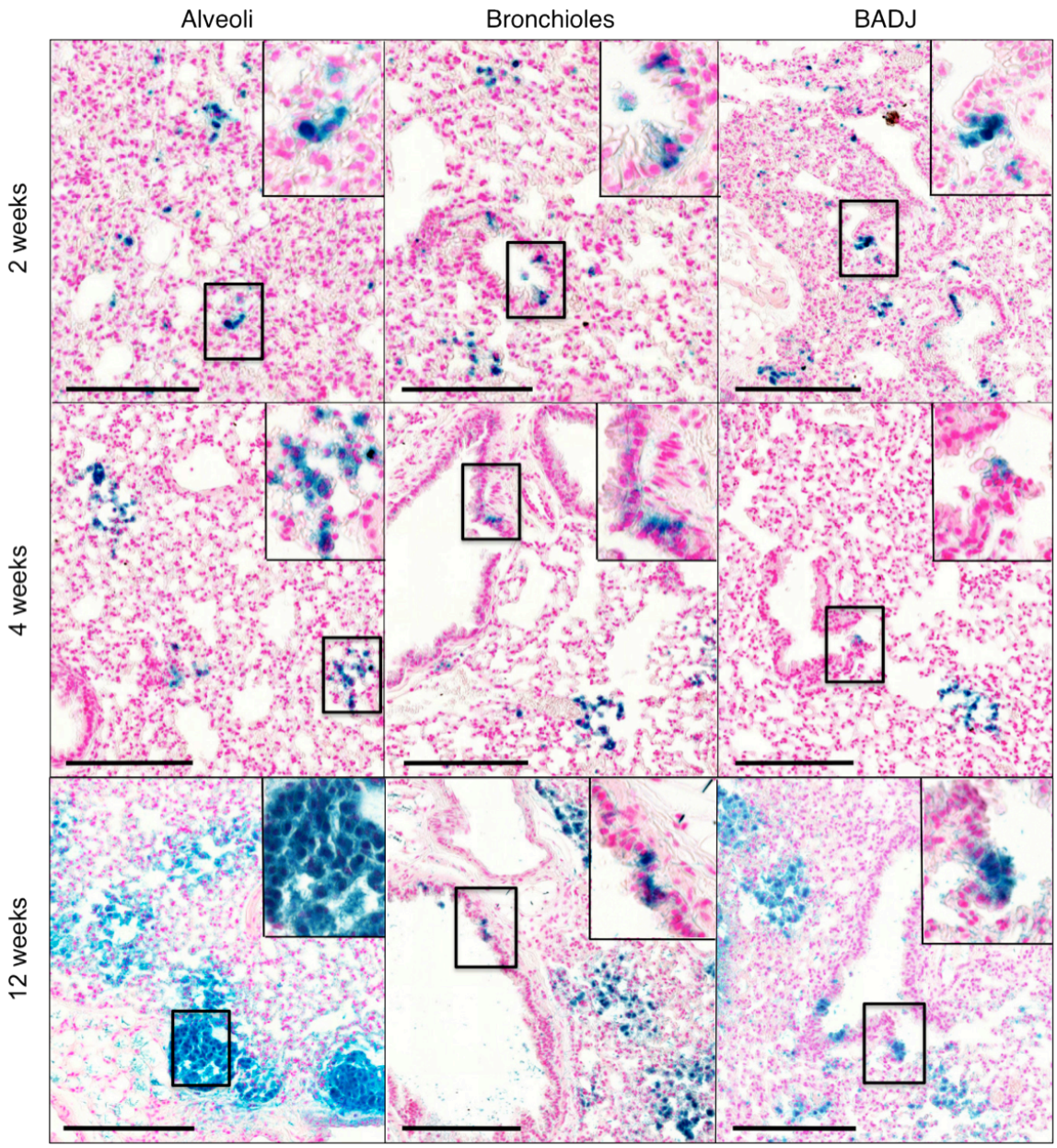


Figure 14: expansion of K-Ras^{G12V}-expressing cells occurs preferentially in the alveoli. K-Ras$^{lox/LSLG12Vgeo}$; RERT$^{ert/ert}$ animals received a single intraperitoneal injection of 4OHT at p21. Mice were sacrificed at the indicated time points (2, 4 and 12 weeks after 4OHT injection) and lungs were submitted to whole mount X-Gal staining. Sections were counterstained with Nuclear Fast Red. Representative pictures are shown. Scale bars represent 200 µm. Inserts: higher magnification.

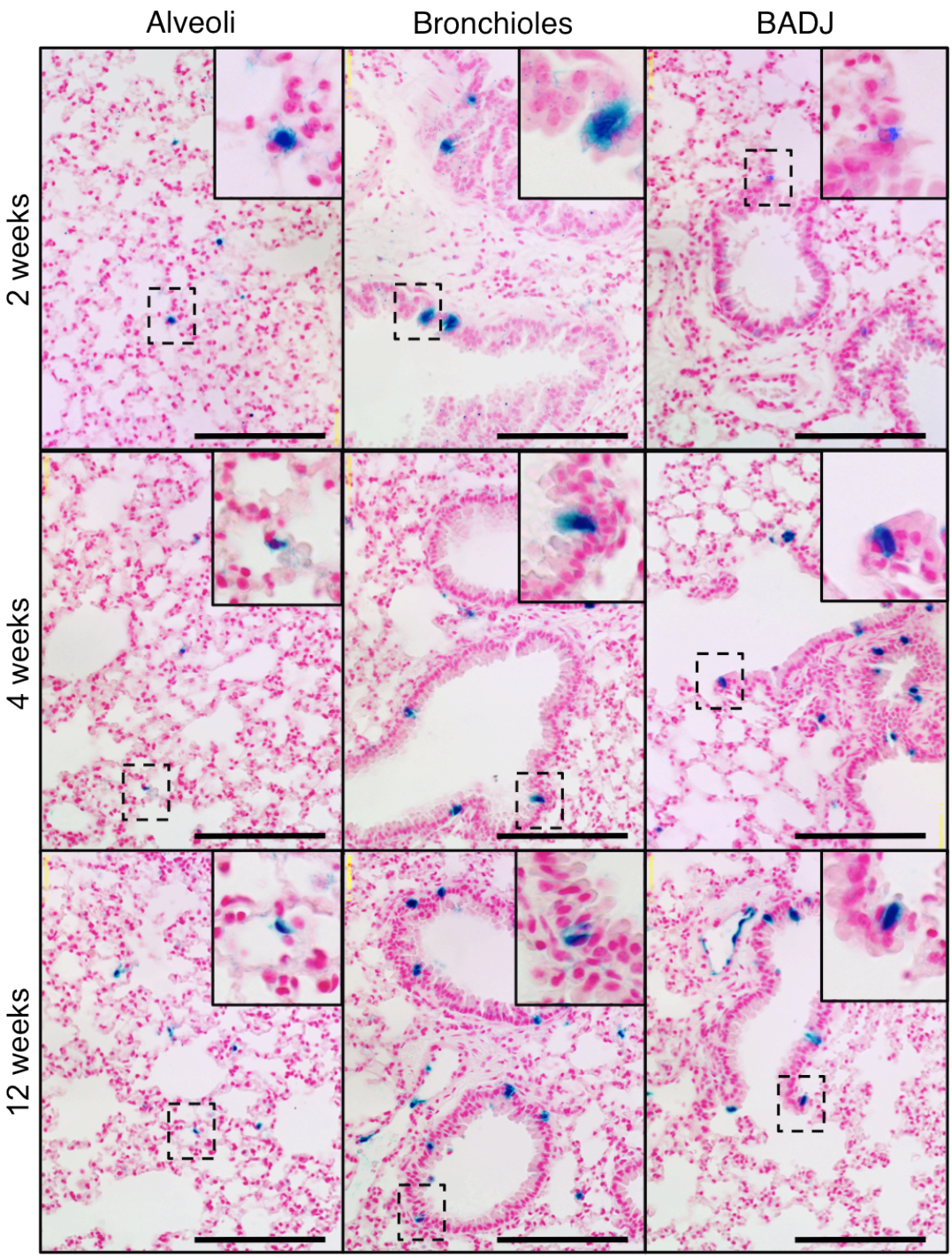


Figure 15: Tracing of β-Gal-expressing cells in reporter mice. Rosa26$^{LSLLacZ/LSLLacZ}$; RERT$^{ert/ert}$ mice animals received a single intraperitoneal injection of 4OHT at p21. Mice were sacrificed at the indicated time points (2, 4 and 12 weeks after 4OHT injection) and lungs were submitted to whole mount X-Gal staining. Sections were counterstained with Nuclear Fast Red. Representative pictures are shown. Scale bars represent 200 µm. Inserts: higher magnification.

When mice were sacrificed 4 weeks post-4OHT exposure, most X-Gal positive cells appeared in clusters, half of which were composed by 4 to 20 cells (measured in 2D sections) whereas the rest had 20 or more cells, including few clusters with as many as 100 cells **(Figure 14 and 16)**. Importantly, whereas clusters containing 4 to 20 cells were uniformly distributed among the main lung structures, those containing more than 20 cells were almost exclusively located in the alveolar region **(Figure 17)**. Moreover, when mice were examined 12 weeks after 4OHT injection, we observed that many X-Gal positive clusters located in the alveoli had expanded into hyperplastic areas **(Figure 14 and 17)**. In contrast, none of the clusters located in either the bronchiolar area or the BADJ region had expanded beyond 10 to 20 cells **(Figure 14 and 17)**. These observations indicate that whereas a large fraction of lung cells can be enticed to divide upon expression of the K-Ras^{G12V} oncoprotein, only cells located in the alveolar region are able to maintain a sustained proliferative response.

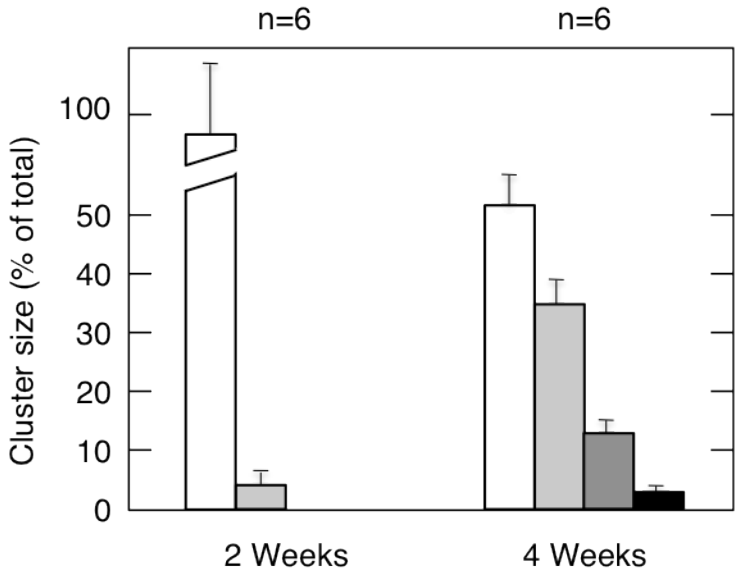


Figure 16: Only a fraction of K-Ras^{G12V}-expressing cells expand beyond 60 cells 4 weeks after 4OHT exposure. Relative abundance of K-Ras^{G12V}-positive clusters according to their size (white bars, 4-20 cells; light grey bars, 21-40 cells; dark grey bar 41-60 cells; black bar >60 cells). K-Ras$^{lox/LSLG12Vgeo}$; RERT$^{ert/ert}$ mice were treated with a single 4OHT injection and sacrificed either 2 or 4 weeks later as indicated. Quantification was performed manually in random sections from β-Gal whole mount stained lungs. n= number of mice.

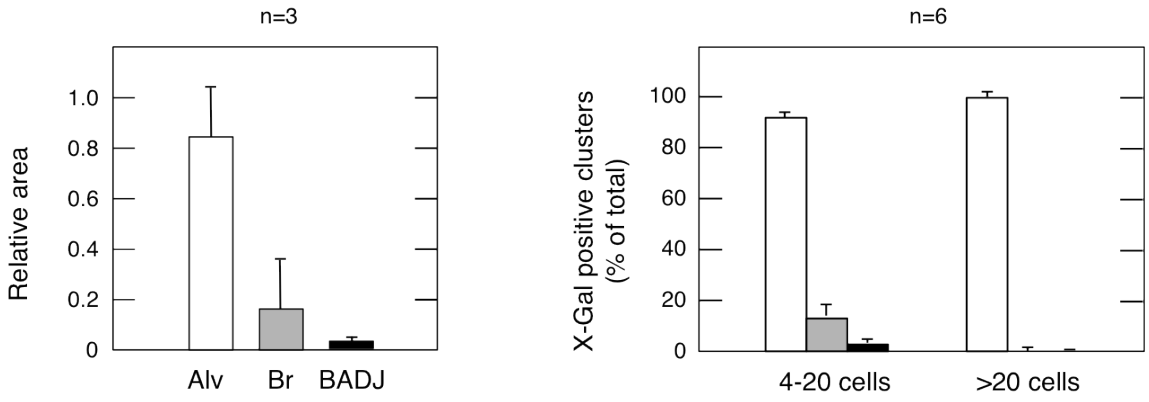


Figure 17: expansion of K-Ras^{G12V}-expressing cells occurs preferentially in the alveoli (quantification). Left panel: quantification of the relative areas occupied in a normal lung by the different types of epithelia: alveoli (Alv, white bar), bronchioles (Br, grey bar), bronchioalveolar duct junctions (BADJs, black bar). Relative areas were estimated using the Panoramic Viewer software (3DHISTECH) in 2D sections from wild type lungs. **Right panel:** Relative abundance of the different K-Ras^{G12V}-positive clusters in different lung regions according to the cluster size (4-20 cells or >20 cells). White bar, alveoli; grey bar, bronchioles; Black bar, BADJs. Counts were performed in 2D sections from K-Ras$^{lox/LSLG12Vgeo}$; RERT$^{ert/ert}$ mice who received a single 4OHT injection and were sacrificed after 12 weeks. n= number of mice.

Lung tissue is made up of few cell types, mainly basal cells, ciliated cells, goblet cells, secretory (Clara) cells (CC), neuroendocrine cells (NE), and alveolar type I (ATI) and type II (ATII) epithelial cells (Rock and Hogan, 2011). A subpopulation of putative progenitor cells, known as BASCs, that express CC and ATII markers have also been identified (Kim et al., 2005). These cell types selectively distribute within the adult lung. Whereas the bronchi primarily contain basal, ciliated and CC cells, the bronchioles are made up of cilliary and CC cells. Both structures also contain NE cells although they are more frequent in the bronchioles. The distal epithelium, or alveoli, is made up of ATI and ATII cells as well as many other cell types of non-lung origin such as macrophages, endothelial cells, pericytes and fibroblasts (Rock and Hogan, 2011). BASCs have been described to specifically reside at the bronchioalveolar duct junctions. In an effort to identify the nature of the cells that respond to K-Ras^{G12V} oncoprotein expression with a sustained proliferative response, we screened adenomas and adenocarcinomas produced by 4OHT-induced activation of K-Ras^{G12V} with markers selective for some of the above cell types including CC10 (also known as Scgb1a/uteroglobin) for secretory

Clara cells, surfactant secretory protein C (SPC/Sftpc) for ATII cells, aquaporin 5 (AQP5) for ATI cells and calcitonin-gene related peptide (CGRP) for NE cells. As illustrated in **Figure 18**, all lesions tested were completely negative for CC10 and CGRP markers. They were also mostly negative for AQP5, except for some positivity at the tumor edges possibly reflecting tumor associated ATI cells. In contrast, all lesions were unambiguously positive for SPC, the marker of ATII cells.

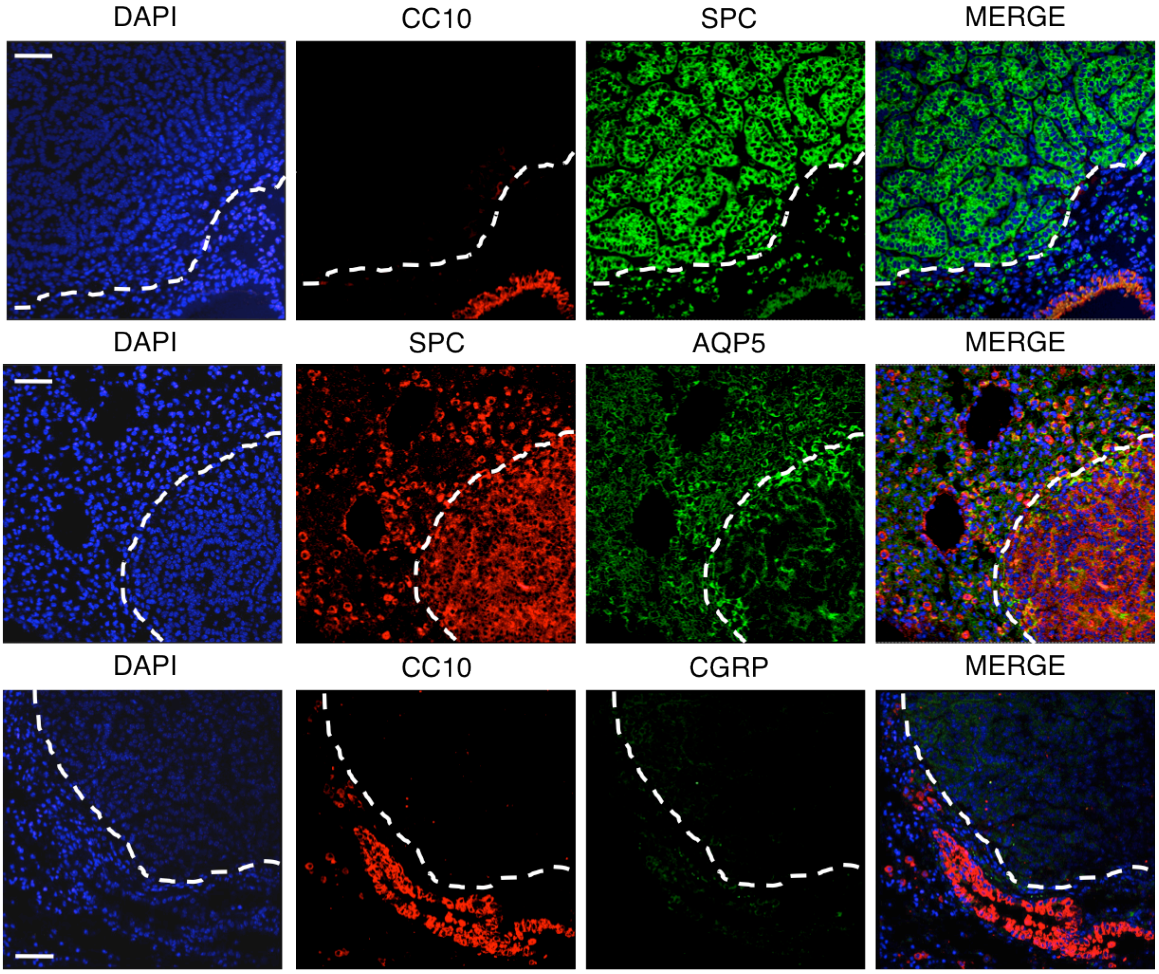


Figure 18: tumoral lesions developed by K-Ras$^{lox/LSLG12Vgeo}$; RERT$^{ert/ert}$ mice are positive for the SPC marker of ATII cells. Immunofluorescence with antibodies against SPC (in green in the upper panel, in red in the middle panel), CC10 (in red in the upper and lower panel), AQP5 (in green in the middle panel) and CGRP (in red in the lower panel) was performed on lung sections from K-Ras$^{lox/LSLG12Vgeo}$; RERT$^{ert/ert}$ mice treated with a single 4OHT injection and sacrificed 8 weeks later. Sections were counterstained with DAPI (in blue). Scale bars represent 50 µm. Dashed lines indicate the tumor edges.

However, previous data regarding the localization of the β-Gal positive clusters suggest that not only ATII cells have an initial proliferative response to K-Ras^{G12V} activation. For this reason we examined mice 4 weeks post 4OHT treatment, when the clusters of K-Ras^{G12V} positive cells (as determined by X-Gal staining) started to proliferate preferentially in the alveolar region. All X-Gal positive clusters containing >20 cells stained for SPC, the marker for ATII cells **(Figure 19)**. None of these clusters were positive for CC10 and AQP5, although we cannot rule out that some X-Gal positive cells may also express AQP5 due to the web-like pattern obtained with AQP5 antibodies **(Figure 19)**. Smaller clusters located in the epithelium lining the bronchioles readily stained for CC10, the marker of Clara cells **(Figure 20)**. Likewise, small clusters of less than 20 X-Gal positive cells present in the BADJ region stained with the CC10 antibodies **(Figure 20)**. A few of these clusters contained cells positive for both CC10 and SPC antibodies suggesting that they correspond to K-Ras^{G12V} expressing BASC cells **(Figure 20)**. None of the cells that stained with CGRP antibodies, the marker for NE cells, were found to co-stain for X-Gal **(Figure 20)**. In summary, all X-Gal positive clusters larger than 20 cells in 2D sections exclusively stained for SPC, the marker of ATII cells **(Figure 21)**.

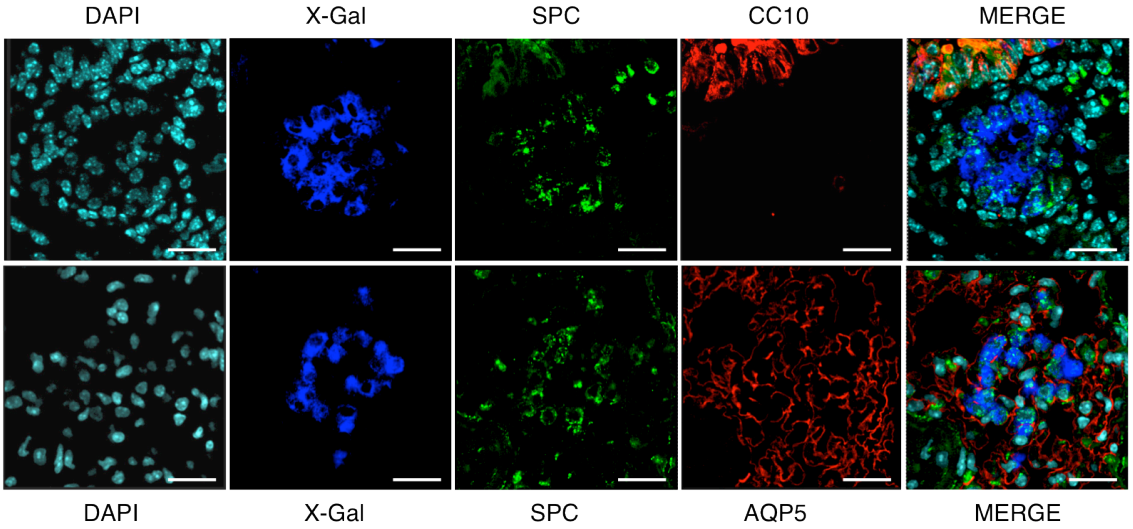


Figure 19: K-Ras^{G12V}-positive clusters 4 weeks after a single 4OHT injection preferentially express the SPC marker of ATII cells. Immunofluorescence with anti-SPC, anti-CC10 and anti-AQP5 antibodies was performed on sections from whole mount X-Gal stained lungs of 4OHT-treated K-Ras$^{lox/LSLG12Vgeo}$; RERT$^{ert/ert}$ mice sacrificed 4 weeks after 4OHT injection. β-Gal-signal was acquired in bright field and overlayed to the confocal IF pictures. Counterstain with DAPI Scale bars represent 20 µm.

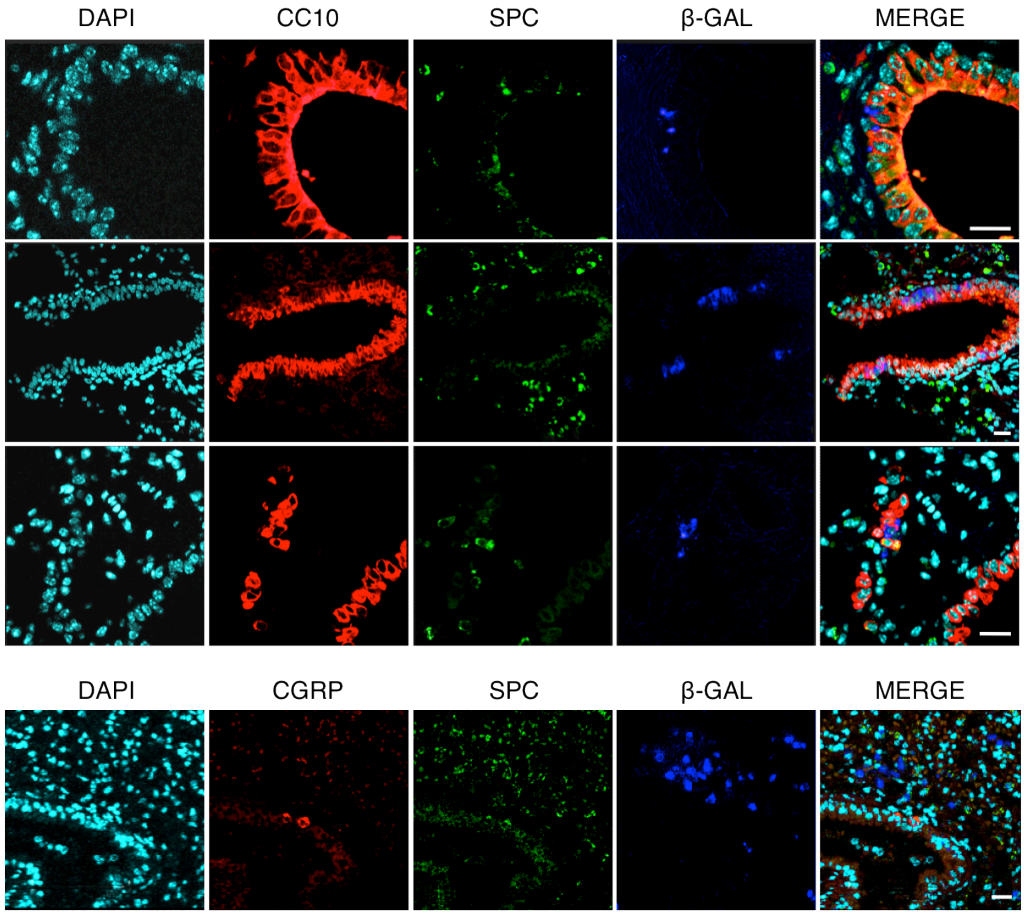


Figure 20: Composition of K-Ras^{G12V}-positive clusters 4 weeks after a single 4OHT injection according to cell type-specific markers expression. Immunofluorescence pictures representative of K-*Ras*^{G12V} expression in different lung cell types. Scale bars represent 20 µm.

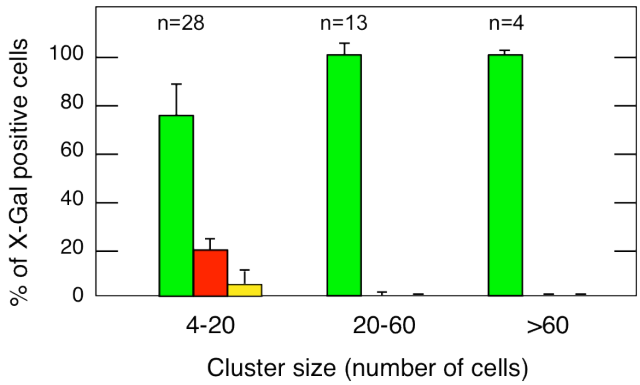


Figure 21: Quantification of K-Ras^{G12V}-positive clusters according to cell type-specific markers expression. K-Ras^{G12V} expressing foci were localized in sections from β-Gal whole mount stained lungs from 2 mice sacrificed 4 weeks after 4OHT injection. For each section scanned in the bright field for the β-Gal signal, the consecutive one was destained from X-Gal precipitates and stained for SPC and CC10 by IF. β-Gal-positive groups were clustered by size and the cell type was assessed in the consecutive section according to the IF. Green bars, percentage of SPC+ cells present in each β-Gal-positive cluster. Red bars, percentage of CC10+ cells. Yellow bars, percentage of double SPC-CC10+ cells (BASCs). n= number of clusters.

6.3 Adult Clara cells can be induced to hyperproliferate by preferential K-Ras^{G12V} expression in the bronchioles.

Although ATII cells are preferentially transformed by the K-Ras^{G12V} oncoprotein following 4OHT administration in K-Ras$^{lox/LSLG12Vgeo}$; RERT$^{ert/ert}$ mice, we observed a peculiar bronchial and bronchiolar hyperplasia in the same mice when infected intratracheally with adeno-Cre virus **(Figure 22 and 24)**. About one third of the tumors that arise 6 months after adeno-Cre infection express both SPC and CC10 markers **(Figure 23 and 24)**. In particular, three different populations coexist in the same K-Ras^{G12V}-induced lesion: ATII, CC and double SPC and CC10 positive cells were all present in the same tumor or bronchiolar hyperplasia **(Figure 24)**. These lesions could derive from the K-Ras^{G12V}-induced proliferation of Clara cells, which could eventually transdifferentiate into SPC or double SPC and CC10 positive cells. Alternatively, the above lesions can originate from bronchioalveolar stem cells (BASCs) present at the bronchioalveolar duct junctions, which express both SPC and CC10 markers and have the intrinsic capacity to give rise to both CC and ATII cells. We reasoned that both technical and biological reasons could underlie the different behavior of adult lung cells when K-Ras^{G12V}-expression is induced either by 4OHT intraperitoneal injection or adeno-Cre intratracheal infection. In fact, on one side, adenoviral delivery through the trachea preferentially leads to activation of the Cre recombinase in the bronchioles, being the first epithelium that physically gets in contact with the virus. On the other side, adenoviral infection evokes an important inflammatory response in the mouse lungs, which attracts inflammatory cells and leads to the production of cytokines and other factors that can eventually influence the proliferation of specific types of adult lung cells. Of note, regardless of the method used for K-Ras^{G12V} activation (either 4OHT or adeno-Cre), a small percentage of the tumors that arise after 6 months do not express neither SPC nor CC10 (1-5%, **Figure 24**), probably as a result of loss of differentiation in very advanced lesions.

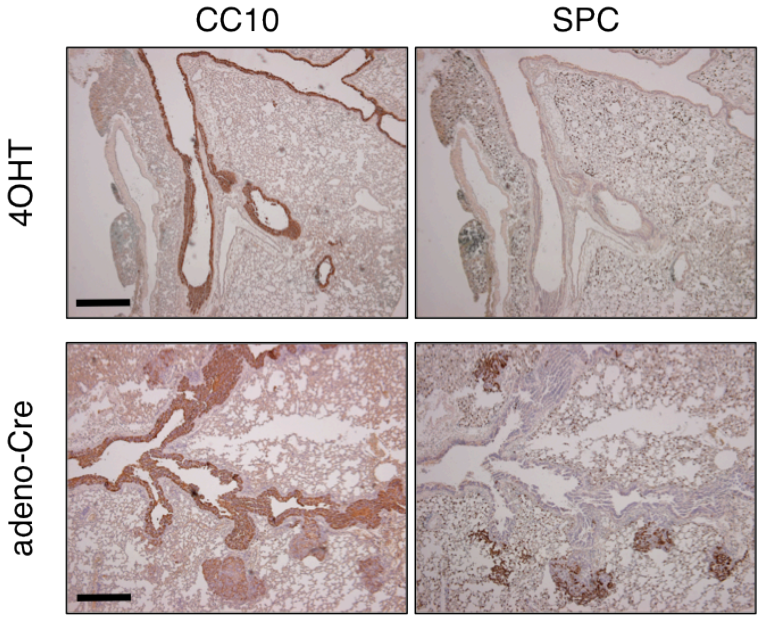


Figure 22: bronchial and bronchiolar hyperplasia in adeno-Cre infected K-Ras$^{lox/LSLG12Vgeo}$ mice. Upper panel, immunohistochemistry showing CC10 and SPC expression in the bronchial tree of 4OHT-treated mice. **Lower panel,** immunohistochemistry showing CC10 and SPC expression in the bronchial tree of adeno-Cre infected mice. The CC10 staining reveals a clear Clara cell hyperplasia while the SPC staining shows the presence of both markers at the bronchioles ends. Scale bars represent 200 µm.

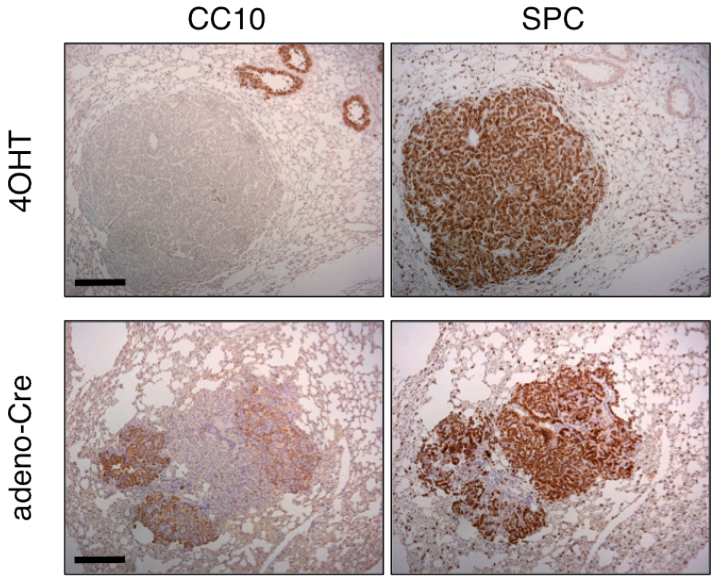


Figure 23: SPC and CC10 positive tumors in adeno-Cre infected K-Ras$^{lox/LSLG12Vgeo}$ mice. Upper panel, immunohistochemistry showing CC10 (negative) and SPC (positive) expression in tumors obtained 6 months after 4OHT treatment. **Lower panel,** immunohistochemistry showing CC10 and SPC expression (both positive) in tumors obtained 6 months after adeno-Cre infection. Scale bars represent 200 µm.

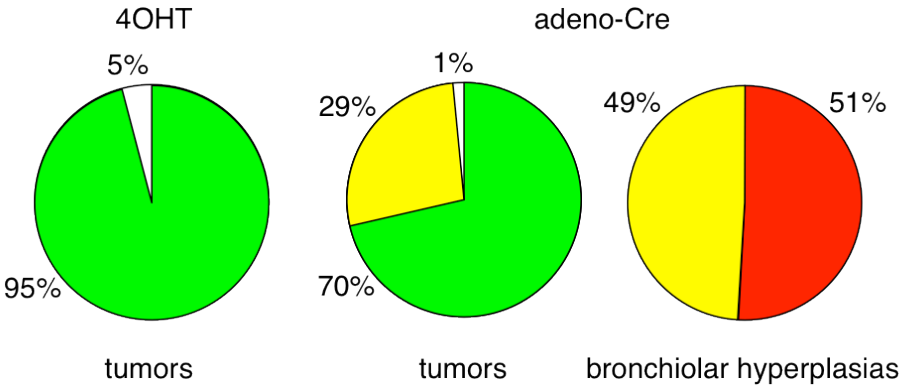


Figure 24: quantification of SPC, CC10 and double positive lesions from either 4OHT (left panel) or adeno-Cre (right panel) treated mice. Quantification was performed 6 months after K-Ras^{G12V} induction on serial sections from 3 different mice per condition. Green area= SPC+, red area= CC10+, yellow area= SPC+CC10+, white area= SPC-CC10-.

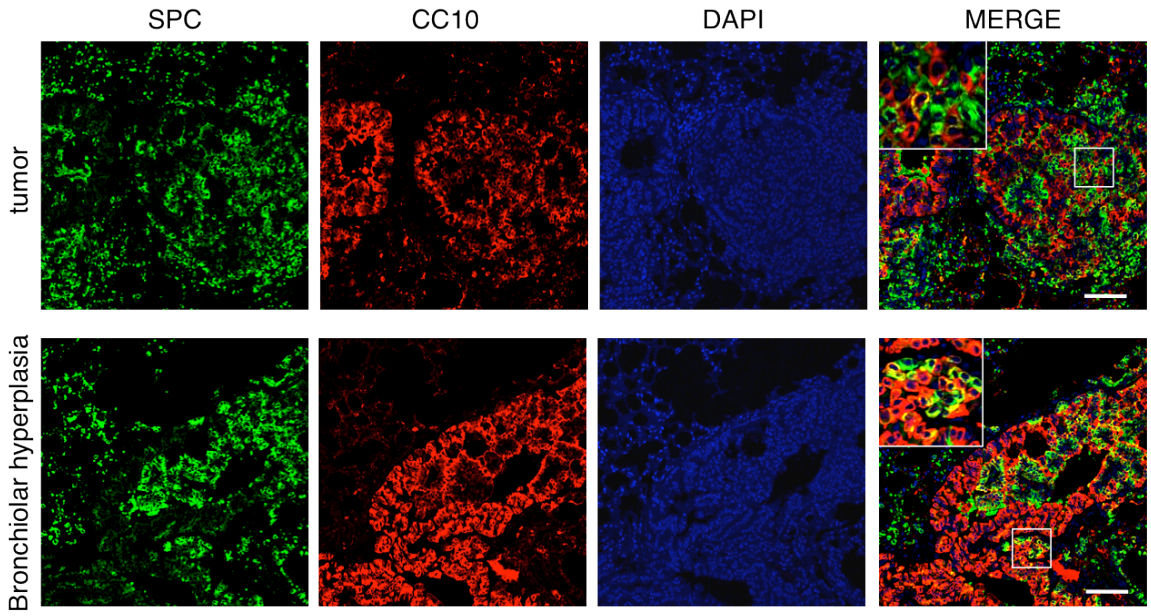


Figure 25: lesions from adeno-Cre infected K-Ras$^{lox/LSLG12Vgeo}$ mice express both SPC and CC10 Representative immunofluorescence pictures of adenomas (upper panel) or bronchiolar hyperplasias (lower panel). Scale bars represent 100 µm. Inserts: higher magnification.

6.4 Developmental timing determines permissiveness to K-Ras^{G12V}-induced transformation.

The above results clearly illustrate that in adult lung an unbiased expression of an endogenous K-Ras^{G12V} oncoprotein, preferentially evokes a sustained proliferative response in ATII cells. To determine whether these observations are due to an intrinsic property of ATII cells, we expressed the resident K-Ras^{G12V} oncogene during embryonic development using a Sca1-Cre transgenic strain known to express the Cre recombinase in most lineages in the embryonic lung due to promoter leakiness (only 50% of the YFP+ cells are actually Sca1+ in a Tg$^{Sca1\text{-}Cre}$; Rosa26$^{+/LSLEYFP}$ reporter mouse, see **Figure 26**). Surprisingly, Tg$^{Sca1\text{-}Cre}$; K-Ras$^{+/LSLG12Vgeo}$ mice rapidly developed lung adenomatous lesions that expressed the CC10 maker (**Figure 28**), thus, suggesting that these tumors originate in secretory Clara cells or their precursors. Interestingly, the Sca1-Cre transgene is also expressed in SPC positive cells during late embryonic development **(Figure 27)**. Ocasionally, we have observed an increased proliferation in these cells leading to the formation of areas with a higher content of ATII cells than that observed in normal tissue (**Figure 29**). Yet, these dense areas never developed into adenomas, at least during the limited life span of these mice.

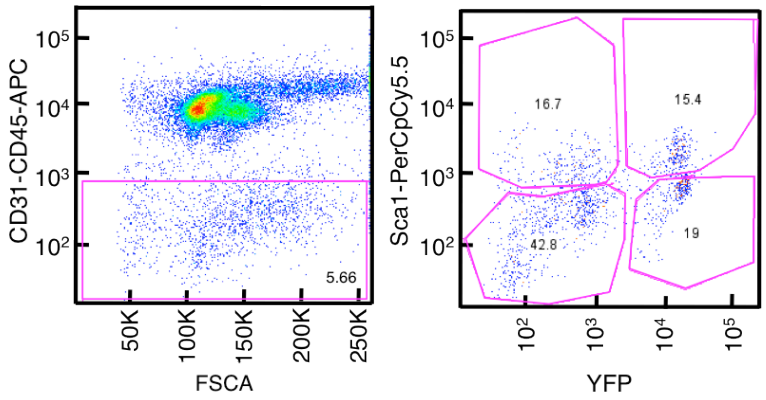


Figure 26: **FACS analysis of Sca1 expression in Tg$^{Sca1\text{-}Cre}$; Rosa26$^{+/LSLEYFP}$ mice, representative experiment.** Mice were sacrificed at 8 weeks of age. Lung single cells preparations were stained with APC-conjugated anti-CD31 and anti-CD45 antibodies to exclude hematopoyetic and endothelial cells, as well as with PerCpCy5.5-conjugated anti-Sca1 antibody**. Left panel**, APC fluorescence (*y* axes) versus forward scatter (*x* axes). The gate is selecting APC negative cells. **Right panel**, PerCpCy5.5 fluorescence (*y* axes) versus YFP. The gates indicate the percentages of YFP- PerCpCy5.5+ (up left), YFP+ PerCpCy5.5+ (up right), YFP- PerCpCy5.5- (low left) and YFP+ PerCpCy5.5- cells (low right).

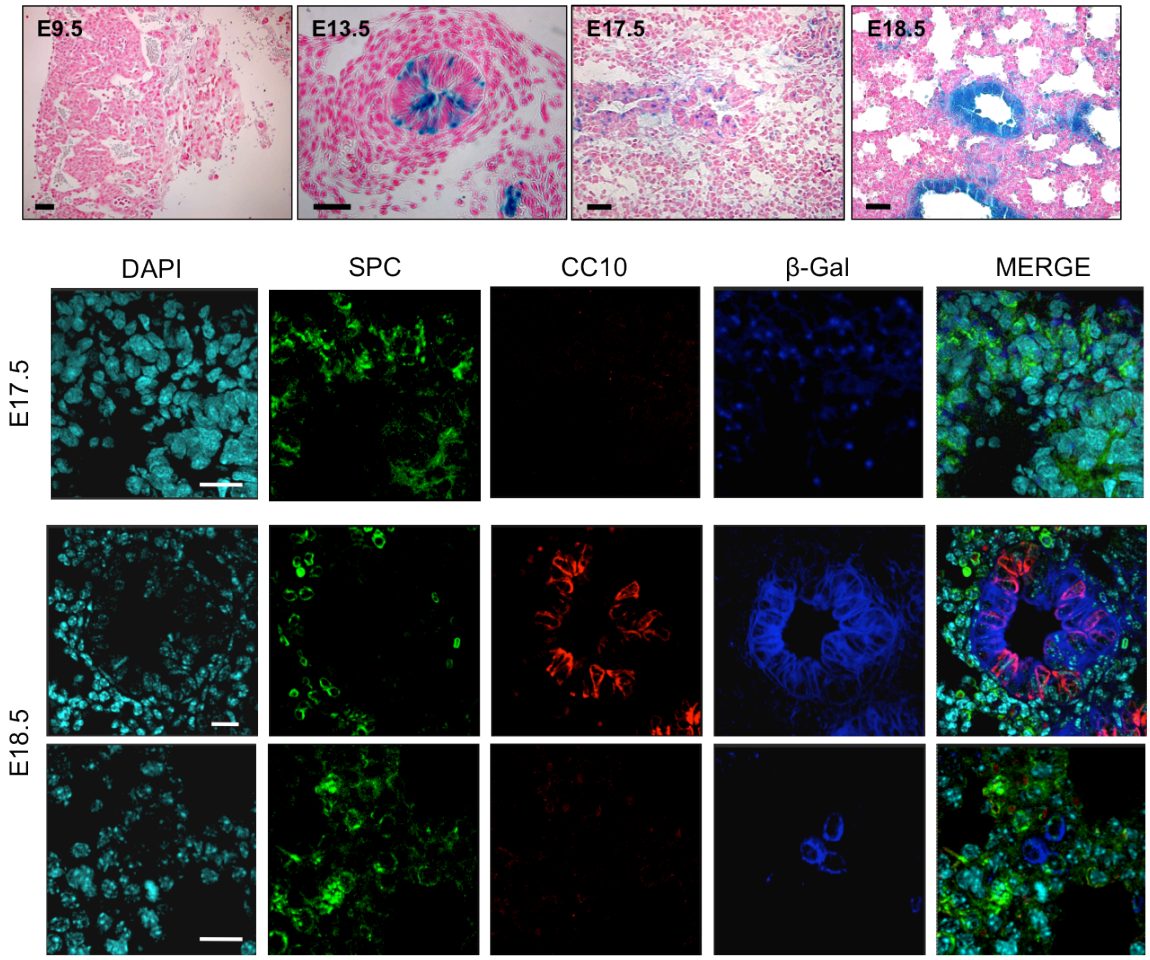


Figure 27: expression of the transgenic Sca1-Cre during lung development. $Tg^{Sca1\text{-}Cre}$; $Rosa26^{+/LSLLacZ}$ embryos were obtained at different times during gestation. **Upper panel,** X-Gal whole mount staining, counterstained with NFR, scale bars represent 50 µm. **Lower panel,** immunofluorescence using antibodies against SPC and CC10. Counterstain with DAPI. Scale bars represent 20 µm.

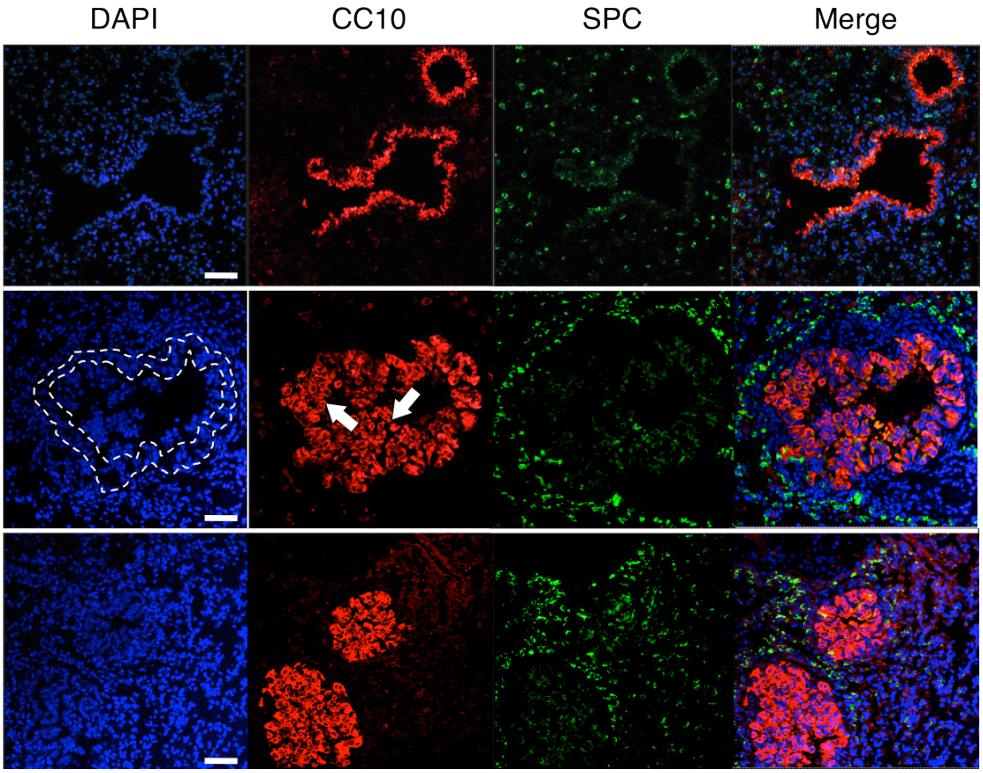


Figure 28: Immunofluorescence showing Clara cell alterations in $Tg^{Sca1\text{-}Cre}$; $K\text{-}Ras^{+/LSLG12Vgeo}$ mice. Upper panel, normal bronchiole and BADJ from control wild type mouse. **Middle panel**, bronchiolar hyperplasia in a section from a 2 months old $Tg^{Sca1\text{-}Cre}$; $K\text{-}Ras^{+/LSLG12Vgeo}$ mouse. The putative original monolayer is marked by dashed line. **Lower panel**, CC10-positive adenomatous lesions in a section from a 2 months old $Tg^{Sca1\text{-}Cre}$; $K\text{-}Ras^{+/LSLG12Vgeo}$ mouse. Scale bars represent 50 µm.

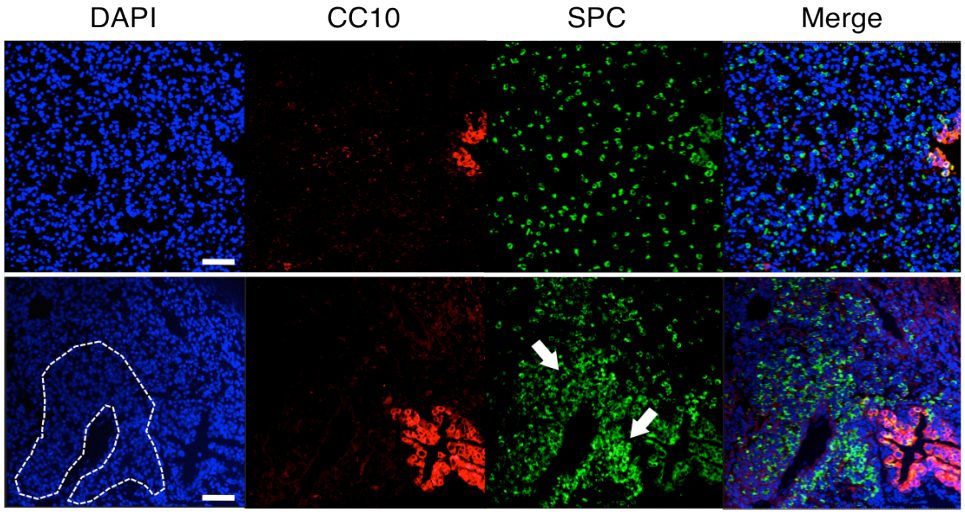


Figure 29: Immunofluorescence showing ATII cells alterations in $Tg^{Sca1\text{-}Cre}$; $K\text{-}Ras^{+/LSLG12Vgeo}$ mice. Upper panel, normal alveoli from a control wild type mouse. **Lower panel**, area of "dense" ATII cells in a 2 months old $Tg^{Sca1\text{-}Cre}$; $K\text{-}Ras^{+/LSLG12Vgeo}$ mouse is marked by dotted line in the DAPI picture and indicated by arrows in the SPC picture. Scale bars represent 50 µm.

We also used a second transgenic strain that expresses an inducible CreERT2 under the control of the same putative Sca1 promoter. X-Gal staining of lung sections derived from $Tg^{Sca1\text{-}CreERT2}$; $Rosa26^{LSLLacZ}$ mice exposed to 4OHT at weaning revealed limited but similar levels of Cre activity in CC10 and SPC positive cells (data not shown). Yet, adult $Tg^{Sca1\text{-}CreERT2}$; K-$Ras^{+/LSLG12Vgeo}$ mice only developed SPC positive lesions **(Figure 30)**. These observations suggest that the resistance of Clara cells to become transformed by K-Ras^{G12V} oncogenes is a property acquired during postnatal development. In addition, they add further support to the hypothesis that ATII cells are the only cell type capable of generating tumor lesions upon activation of K-Ras oncogenes in adult mice.

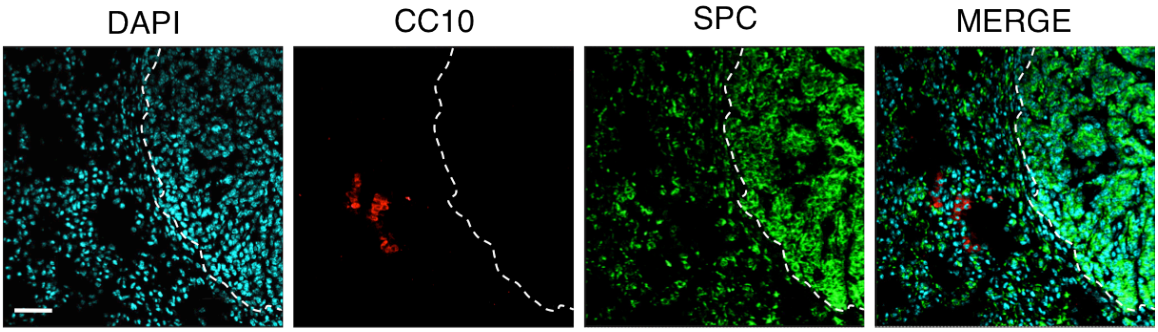


Figure 30: tumors arised in $Tg^{Sca1\text{-}CreERT2}$; K-$Ras^{+/LSLG12Vgeo}$ mice express the SPC marker of ATII cells. Immunofluorescence with antibodies against SPC and CC10. Mice were treated with 4OHT and sacrificed after 2 months. Tumor edge is marked by the dashed line. Scale bar represents 50 µm.

6.5 Further characterization of $Tg^{Sca1\text{-}Cre}$; K-$Ras^{+/LSLG12Vgeo}$ mice

$Tg^{Sca1\text{-}Cre}$; K-$Ras^{+/LSLG12Vgeo}$ mice are not born at a mendelian ratio **(Table 1)**, meaning that expression of the K-Ras oncogene during embryogenesis affects the embryonic development although with incomplete penetrance. The $Tg^{Sca1\text{-}Cre}$; K-$Ras^{+/LSLG12Vgeo}$ mice that survive usually die around 12 weeks of age due to multiple defects. Macroscopically, those mice are smaller than the wild type littermates, they show hair loss and craniofacial abnormalities **(Figure 31)**. Also at the histological level, $Tg^{Sca1\text{-}Cre}$; K-$Ras^{+/LSLG12Vgeo}$ mice present defects in various organs **(Figure 32)**. In particular, in the skin

we could observe an increased accumulation of melanin, hyperkeratinization, atrophied hair follicles, reduced thickness of the epidermal layer and excess of nerves in the dermis. In the stomach we could often observe papillomas, characterized by the presence of lots of mitotic figures, as well as keratin accumulation and cystic glandules. The spleen was atrophic and disorganized, due to almost complete loss of the red pulp. In the skeletal muscle, we could observe the same nerve hypertrophy that we found in the skin. The pancreas showed acinar atrophy, characterized by loss of acinar granules, autophagy vacuoles and reduced number of pancreatic islets. The liver was hyperplastic and presented several foci of hepatitis. At least some of these phenotypes could be associated to a disfunction in a progenitor population (being Sca1 a general marker of stem cells). For example excessive proliferation (due to K-Ras^{G12V} activation) of skin stem cells during development could lead to exhaustion of the stem compartment and subsequent hair loss.

$Tg^{Sca1\text{-}Cre/+}$ K-$Ras^{+/LSLG12Vgeo}$	**$Tg^{Sca1\text{-}Cre/+}$ K-$Ras^{+/+}$**	**$Tg^{+/+}$ K-$Ras^{+/+}$**	**$Tg^{+/+}$ K-$Ras^{+/LSLG12Vgeo}$**
14%	29%	33%	24%

Table1: percentages of the different genotypes obtained in P21 pups from $Tg^{Sca1\text{-}Cre/+}$; K-$Ras^{+/+}$ X $Tg^{+/+}$; K-$Ras^{+/LSLG12Vgeo}$ crosses (expected ratio 25% each).

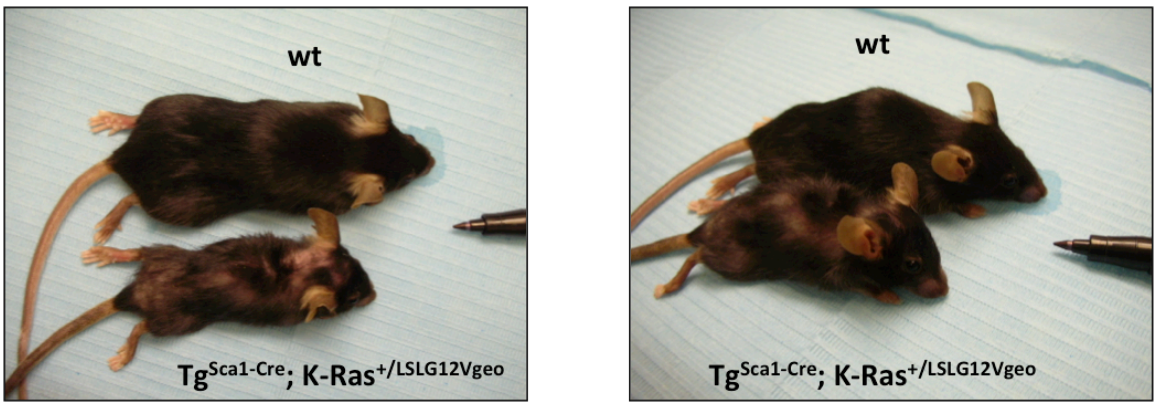


Figure 31: macroscopic phenotype of 2 littermates 8 weeks old $Tg^{Sca1\text{-}Cre}$; K-$Ras^{+/LSLG12Vgeo}$ male mice.

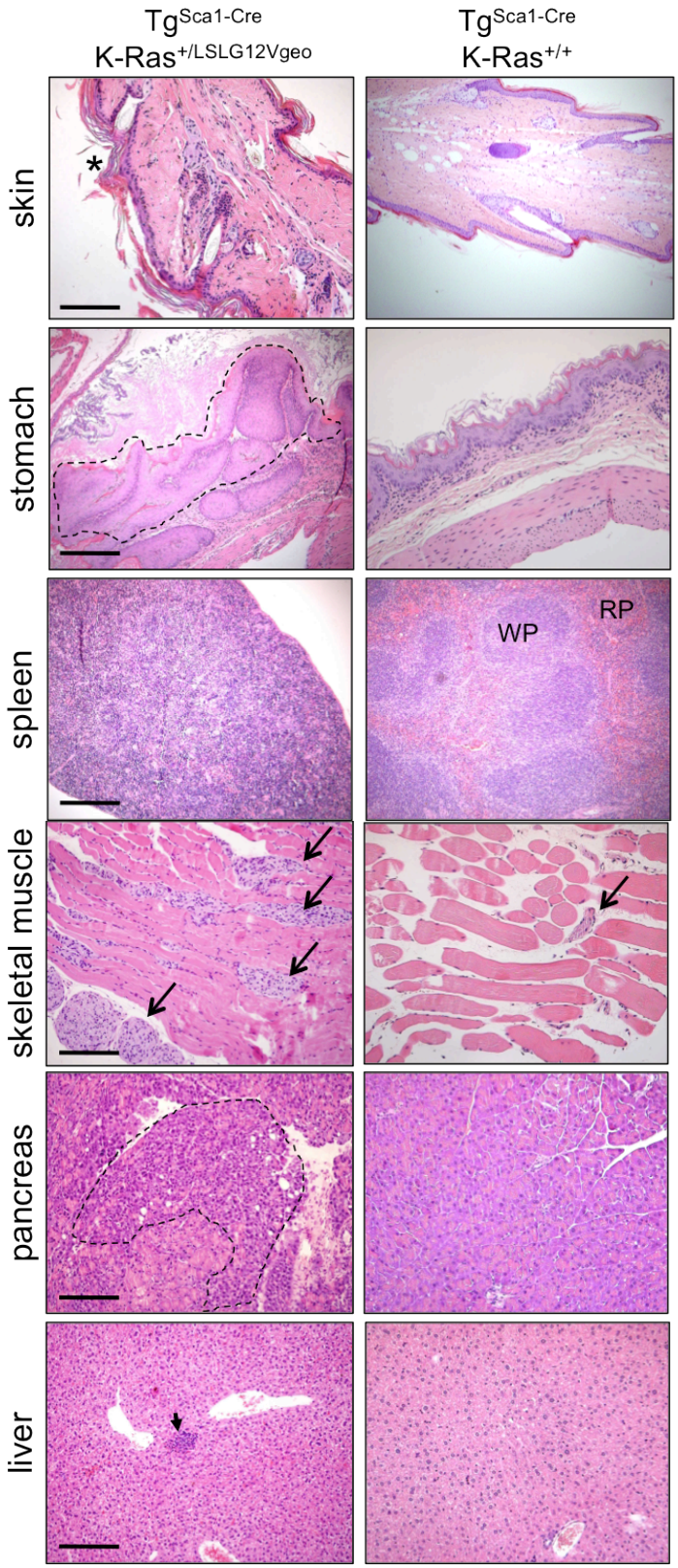


Figure 32: histological analysis of $Tg^{Sca1\text{-}Cre}$; $K\text{-}Ras^{+/LSLG12Vgeo}$ mice. H&E stainings of the indicated organs. Scale bars represent 100 µm. Asterisk: hyperkeratinization of the skin. Dashed line in stomach: papilloma. WP: white pulp. RP: red pulp. Arrows: nerves. Dashed line in pancreas: area of degranultated acini. Arrowhead: focus of hepathitis.

6.6 Transcriptome alterations during the earliest stages of tumor development

In order to gain insights on the molecular events accompanying the very first steps of K-Ras^{G12V}-induced proliferation, we performed a gene expression analysis of K-Ras^{G12V}-expressing lung cells at early time points. To this purpose, we crossed K-Ras$^{+/LSLG12Vgeo}$ mice with a reporter Rosa26$^{LSLEYFP/LSLEYFP}$ strain, in order to introduce the YFP protein as a fluorescent marker. K-Ras$^{+/LSLG12Vgeo}$; Rosa26$^{LSLEYFP/LSLEYFP}$ mice were subjected to adeno-Cre intratracheal infection in order to produce a massive and localized expression of an active Cre recombinase. Cells expressing YFP one month after infection are supposed to express an active Cre and, concomitantly, the oncogenic form of K-Ras protein in the K-Ras$^{+/LSLG12Vgeo}$; Rosa26$^{LSLEYFP/LSLEYFP}$ genotype. Thus, YFP marker can be used to isolate K-Ras^{G12V}-expressing cells through fluorescence-activated cell-sorting techniques **(Figure 33)**. 4 weeks after adeno-Cre intratracheal infection, 5% of lung epithelial cells express the YFP protein in K-Ras$^{+/+}$; Rosa26$^{LSLEYFP/LSLEYFP}$ mice, while the YFP+ population increases to 30% in K-Ras$^{+/LSLG12Vgeo}$; Rosa26$^{LSLEYFP/LSLEYFP}$ mice due to the expression of K-Ras oncoprotein **(Figure 33)**. Single cells preparations were obtained from 4 K-Ras$^{+/LSLG12Vgeo}$; Rosa26$^{LSLEYFP/LSLEYFP}$ and 4 control K-Ras$^{+/+}$; Rosa26$^{LSLEYFP/LSLEYFP}$ mice 4 weeks after adeno-Cre intratracheal infection. CD31-CD45-negative (non hematopoietic, non endothelial), YFP-positive cells were sorted and RNA isolated and hybridized on Agilent mouse 60K chip. Analysis of the transcriptomes using an FDR value <0.01 revealed more than 900 significant differentially regulated genes **(Figure 34, Table 2 and 3)**, demonstrating that as early as one month after the oncogene expression, when proliferative clusters are still very small and morphologically indistinguishable from normal tissue, the expanded cell population has already activated a specific and complex transcriptional program.

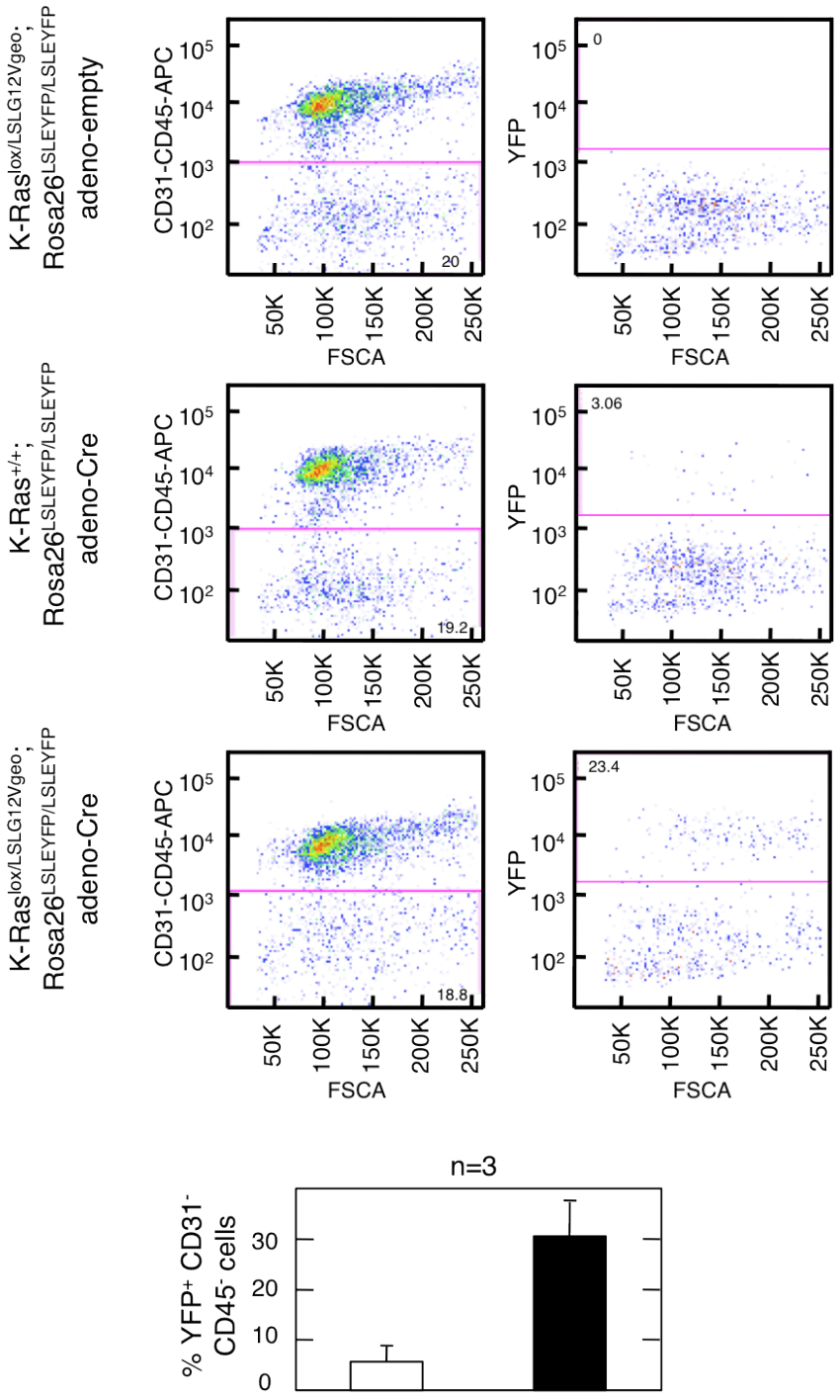


Figure 33: isolation of early K-Ras^{G12V}-expressing YFP-positive cells. Upper panel, representative sorting experiment. Lung cells preparation from mice with the indicated genotypes infected with the indicated adenovirus and sacrificed 4 weeks after the infection. **Lower panel,** quantification of CD31-CD45-negative YFP-positive population in K-Ras$^{+/+}$; Rosa26$^{LSLEYFP/LSLEYFP}$ (white bar) and K-Ras$^{+/LSLG12Vgeo}$; Rosa26$^{LSLEYFP/LSLEYFP}$ (black bar) mice 4 weeks after Adeno-Cre intratracheal infection. n= number of mice.

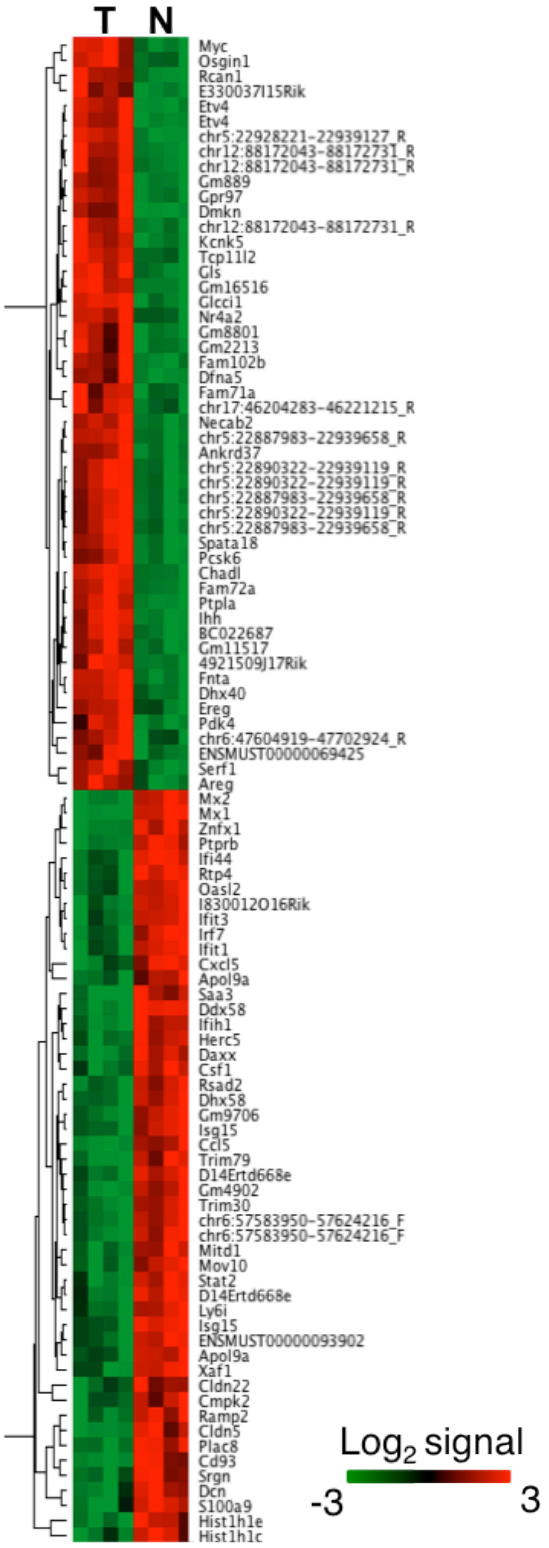


Figure 34: gene expression analysis of early proliferating K-Ras^{G12V}-expressing cells. Heat map of the 100 best differentially expressed genes from the Agilent Mouse 60K microarray: T= K-Ras$^{lox/LSLG12Vgeo}$; Rosa26$^{LSLEYFP/LSLEYFP}$ CD31-CD45-negative YFP-positive cells; N= K-Ras$^{+/+}$; Rosa26$^{LSLEYFP/LSLEYFP}$ CD31-CD45-negative YFP-positive cells.

Accession number	Gene symbol	Adjusted P-value	Fold change	Function
NM_008815	Etv4	6.68 E-05	6.80	DNA binding, regulation of branching involved in organ morphogenesis, stem cells differentiation
NM_054095	Necab2	7.30 E-05	6.43	Calcium ion binding
chr12:88172043-88172731_R		1.08 E-04	6.55	NA
NM_010849	Myc	2.26 E-04	5.39	DNA binding, regulation of cell proliferation, growth, apoptosis, differentiation, stem cells self-renewal
chr5:22890322-22939119_R		2.72 E-04	4.66	NA

Table 2. The 5 most significant differentially upregulated genes in K-Ras$^{lox/LSLG12Vgeo}$; Rosa26$^{LSLEYFP/LSLEYFP}$ cells *versus* K-Ras$^{+/+}$; Rosa26$^{LSLEYFP/LSLEYFP}$. NA = no information available.

Accession number	Gene symbol	Adjusted P-value	Fold change	Function
NM_013606	Mx2	1.01 E-05	-23.25	GTPase activity, cytokine-mediated signaling pathway
NM_010846	Mx1	1.01 E-05	-14.26	GTPase activity, innate immune response
NM_013653	Ccl5	4.02 E-05	-279.57	Chemokine receptor binding
NM_001033196	Znfx1	4.72 E-05	-6.47	Metal ion binding
NM_001005858	I830012O16Rik	6.68 E-05	-34.47	NA

Table 3. The 5 most significant differentially downregulated genes in K-Ras$^{lox/LSLG12Vgeo}$; Rosa26$^{LSLEYFP/LSLEYFP}$ cells *versus* K-Ras$^{+/+}$; Rosa26$^{LSLEYFP/LSLEYFP}$. NA= no information available.

Of note, the proto-oncogene Myc appeared to be the most upregulated gene in K-Ras^{G12V} positive sorted cells, providing a validation of the reliability of our approach to isolate early transformed cells, as it was already demonstrated that Myc is essential for mutant K-Ras induced lung adenocarcinomas (Soucek et al., 2008).

Interestingly, the transcription factor Etv4 was aslo found among the most upregulated genes in early K-Ras^{G12V}-expressing cells. Etv4 is a member of the Ets transcription factor superfamily, characterized by highly related DNA-binding domains containing winged-helix-turn-helix motifs. Among these factors, a subfamily of highly related Ets proteins comprising Etv1, Etv4 and Etv5 has been described as playing an important role during lung development. Moreover, Etv4 overexpression has been described in NSCLC (Hiroumi et al., 2001). In order to confirm Etv4 upregulation during different stages of lung tumor initiation and progression, we performed quantitative RT PCR on lung samples from K-Ras$^{lox/LSLG12Vgeo}$; Rosa26$^{LSLEYFP/LSLEYFP}$ mice sorted either 1 month after adeno-Cre infection or 2 months after a single 4OHT injection, as well as from microdissected lesions 3 months after adeno-Cre infection or fully grown tumors isolated 6 months after adeno-Cre infection. In all cases we observed an increase in Etv4 mRNA levels as compared to controls, with the highest upregulation observed in 6 months tumors **(Figure 35)**. Moreover, Etv4 overexpression could also be detected at the protein level in K-Ras^{G12V}-driven lung tumors, either by western blot or by immunohistochemistry **(Figure 35)**.

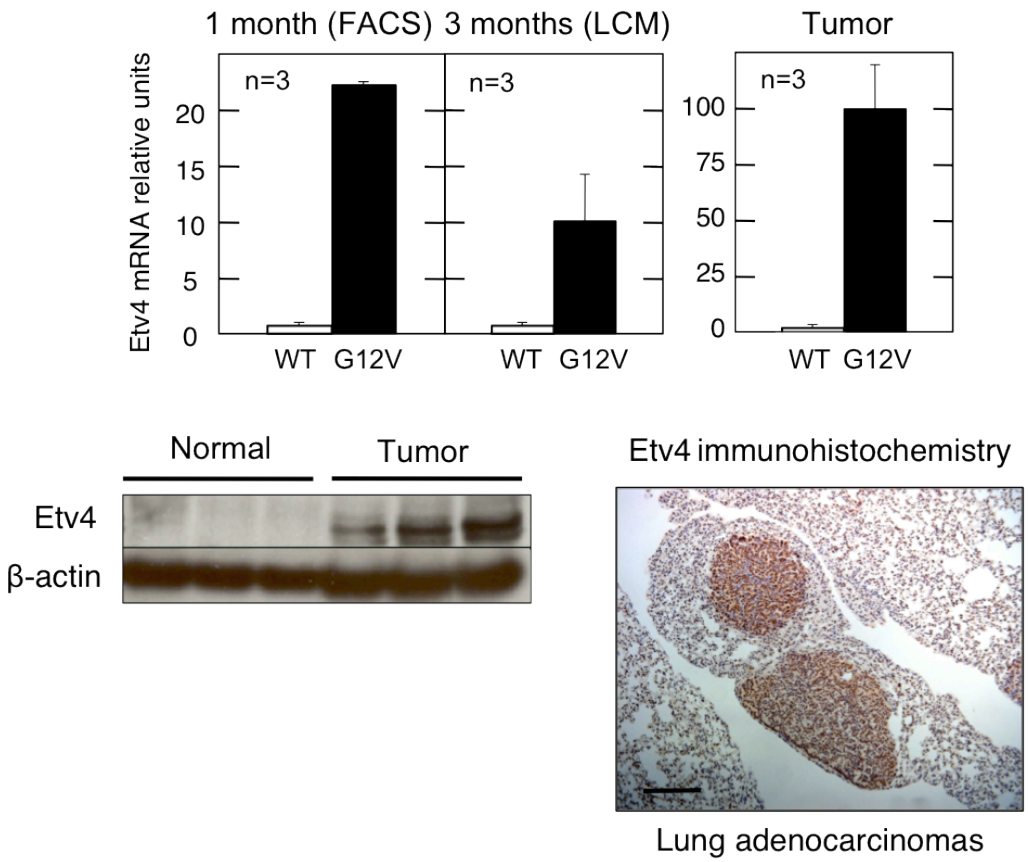


Figure 35: Etv4 is overexpressed in K-RasG12V-induced lung tumors and early lesions. Upper panel, quantitative RT PCR on Etv4 mRNA. Left: white bar, sorted K-Ras$^{+/+}$; Rosa26$^{LSLEYFP/LSLEYFP}$ CD31-CD45-negative YFP-positive cells from adeno-Cre infected mice sacrificed at 72 hs postinfection; black bar, sorted K-Ras$^{lox/LSLG12Vgeo}$; Rosa26$^{LSLEYFP/LSLEYFP}$ CD31-CD45-negative YFP-positive cells from adeno-Cre infected mice sacrificed 4 weeks postinfection. Middle: white bar, microdissected lung alveolar regions from adeno-GFP infected K-Ras$^{lox/LSLG12Vgeo}$ mice; black bar, microdissected lesions from adeno-Cre infected K-Ras$^{lox/LSLG12Vgeo}$ mice. Both groups of mice were sacrificed 3 months after infection. Right: white bar, wild type lung; black bar, tumors dissected from K-Ras$^{lox/LSLG12Vgeo}$ mice 6 months after adeno-Cre infection**. Lower left panel:** western blot analysis of Etv4 protein. Single tumors were dissected from K-Ras$^{lox/LSLG12Vgeo}$ mice 6 months after adeno-Cre infection. **Lower right panel:** immunohistochemistry using anti-Etv4 antibody. Representative section from K-Ras$^{lox/LSLG12Vgeo}$; RERT$^{ert/ert}$ mouse treated at weaning with 4OHT and sacrificed 6 months after treatment. Scale bar represents 200 µm.

6.6 Etv4 is important, but not essential, for K-Ras^{G12V}-induced lung tumor initiation.

In order to study the role of Etv4 in adult lung and lung tumor initiation, we took advantage of Etv4$^{+/NLacZ}$ mice, which have a bacterial β-Galactosidase expression cassette "knocked-in" into the endogenous Etv4 locus, and express β-Gal in a pattern that mimics the spatial and temporal expression of the endogenous Etv4 allele. Adult Etv4$^{+/NLacZ}$ mice express the β-Gal reporter in a very limited number of cells in the lung **(Figure 36)**: in proximal airways, namely trachea and major bronchi, it colocalizes with the CK5 marker of basal cells **(Figure 37)**; in distal airways β-Gal signal does not colocalize neither with SPC, CC10, AQPV or CGRP markers, and it is found in very few peribronchiolar cells of unknown origin **(Figure 38)**. Following the observation that Etv4 is upregulated in the early K-Ras^{G12V}-expressing and proliferating cells we decided to cross our NSCLC model with the Etv4^{NLacZ} mice above described. K-Ras$^{+/FSFG12V}$ mice (in which the oncogene expression can be activated by FlpO recombinase and which lacks the β-Gal reporter) were used to set up crosses instead of the KRas$^{+/LSLG12Vgeo}$ line, in order to avoid overlapping between two β-galactosidase reporters. K-Ras$^{+/FSFG12V}$; Etv4$^{+/NLacZ}$ mice were infected with a high titer (6x10^{7} pfu/mouse) of adeno-FlpO virus at 8 weeks of age. One week after infection β-Gal positive Etv4-expressing cells started to be detectable in the alveolar region and (more often) in the bronchioles **(Figure 39).** As described above, Etv4 is not normally expressed in alveolar or bronchiolar cells, leading to the conclusion that K-Ras is able to induce *de novo* Etv4 expression in these cell types. One month after adeno-FlpO infection very small and isolated clusters (3-4 cells) of Etv4-positive cells could be detected in both bronchiolar and alveolar epithelia **(Figure 39)**. Surprisingly, 3 months after infection K-Ras$^{+/FSFG12V}$; Etv4$^{+/NLacZ}$ mice had developed much less pretumoral lesions compared to the control K-Ras$^{+/FSFG12V}$; Etv4$^{+/+}$ mice **(Figure 40)**. The few adenomas that arose in K-Ras$^{+/FSFG12V}$; Etv4$^{+/NLacZ}$ mice were positive for Etv4, as assessed by X-Gal staining **(Figure 40)**. 6 months after adeno-FlpO infection K-Ras$^{+/FSFG12V}$; Etv4$^{+/NLacZ}$ mice still show a smaller number of tumors as compaed to the K-Ras$^{+/FSFG12V}$; Etv4$^{+/+}$ controls. No obvious difference could be appreciated in the grade or

size of the adenomas developed by the two genotypes. Altogether, these results clearly indicate that Etv4 plays an important role in K-Ras^{G12V}-induced lung tumor initiation, and that ablation of a single Etv4 allele already affects tumor formation. Despite this, Etv4$^{NLacZ/NLacZ}$ mice, which lack expression of Etv4, still present a number of lesions similar to the Etv4$^{+/NLacZ}$ mice3 months after adeno-FlpO infection, indicating that Etv4 is not totally essential for K-Ras^{G12V}-driven lung tumor initiation. We hypothesized that in absence of Etv4, other members of the Ets family, possibly Etv1 and Etv5, which display high homology with Etv4, may compensate for its lack and activate its downstream targets. To this purpose we analyzed by RT Real Time PCR the mRNA levels of the three related genes **(Figure 41)** in tumors isolated 6 months after adeno-FlpO infection from either K-Ras$^{+/FSFG12V}$; Etv4$^{+/+}$ or K-Ras$^{+/FSFG12V}$; Etv4$^{+/NLacZ}$ (Etv4 KO tumors could not be analyzed because mice from this genotype did not reach the 6 months time point yet). Accordingly with Etv4 being present as only one allele, its mRNA levels appeared reduced at about one half in the K-Ras$^{+/FSFG12V}$; Etv4$^{+/NLacZ}$ tumors. Unexpectedly, Etv1 mRNA levels were also downregulated. On the contrary the mRNA levels of Etv5 were significantly upregulated in K-Ras$^{+/FSFG12V}$; Etv4$^{+/NLacZ}$ tumors. Thus, in a situation in which Etv4 expression is reduced, compensatory Etv5 upregulation may work to maintain the levels of Ets target genes at a level sufficient to promote tumor growth.

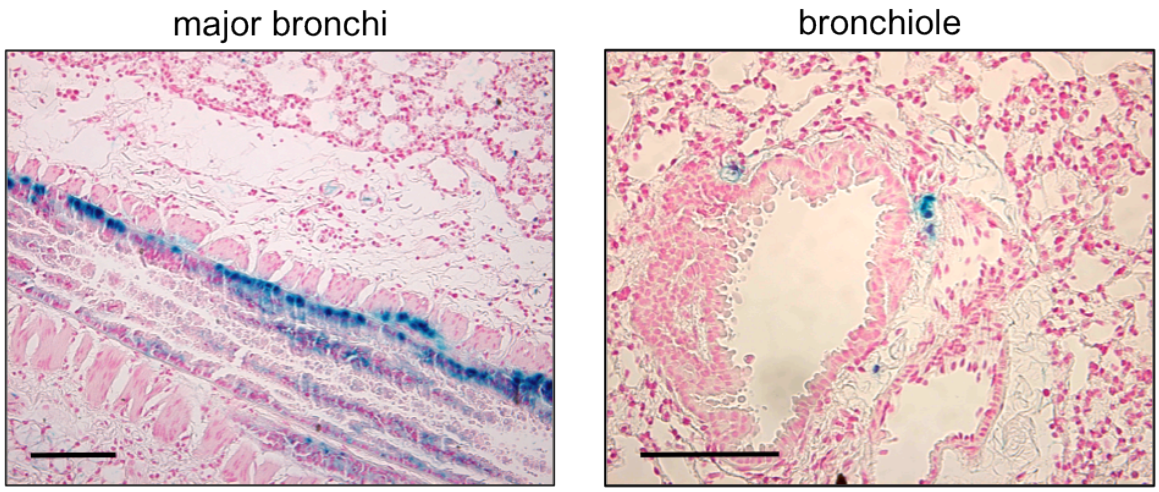


Figure 36: Etv4 in expressed in few isolated cells in normal adult lung. Representative sections from X-Gal whole mount stained lung of an $Etv4^{+/NLacZ}$ mouse, showing few Etv4-positive cells either in a basal localization (in major bronchi, on the left) or in a peribronchiolar localization (on the right). The section is counterstained with NFR. Scale bars represent 100 µm.

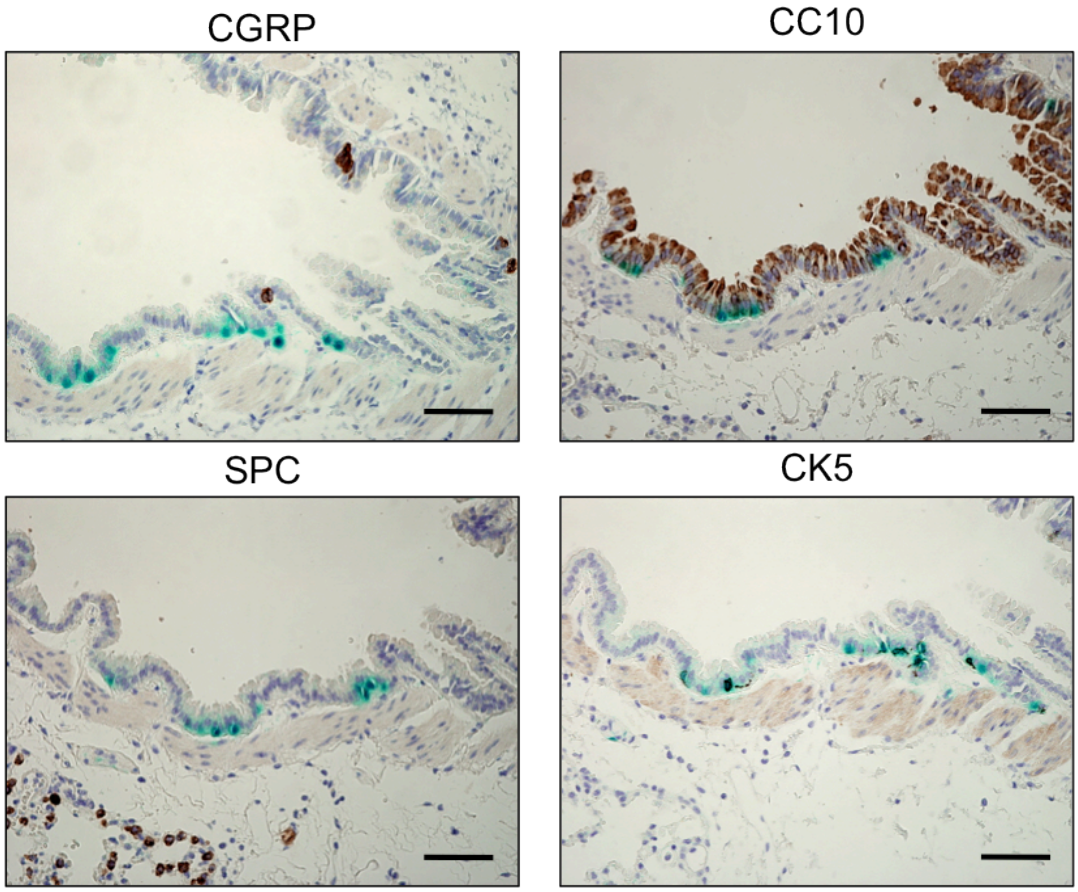


Figure 37: Etv4 is expressed in basal cells in normal adult lung. CGRP**,** CC10, SPC, and CK5 immunohistochemistry was performed on consecutive sections from X-Gal whole mount stained lungs of $Etv4^{+/NLacZ}$ mice. Scale bars represent 50 µm

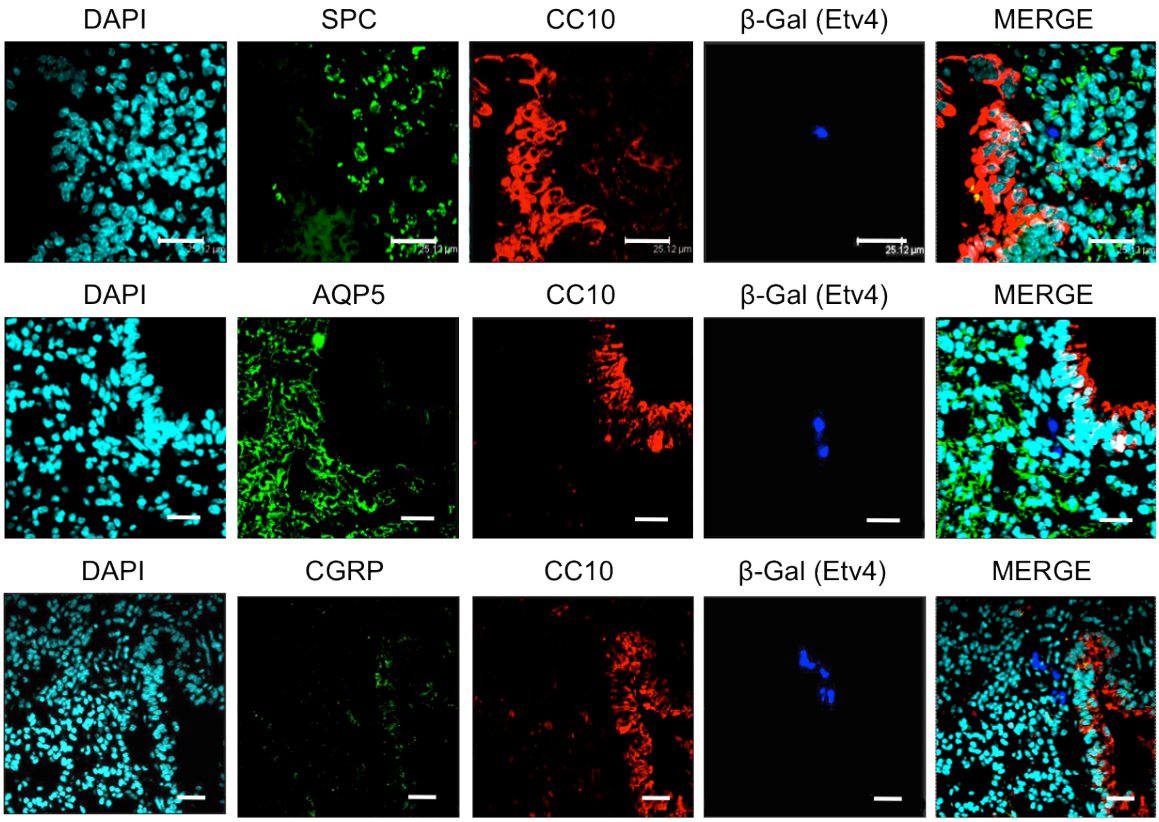


Figure 38: peribronchiolar Etv4 does not colocalize with the most common lung cell type-specific markers. SPC, CC10, AQP5 and CGRP immunofluorescence was performed on sections from X-Gal whole mount stained lungs of Etv4$^{+/NLacZ}$ mice in which the blue staining represents Etv4 expressing cells. Scale bars represent 25 µm.

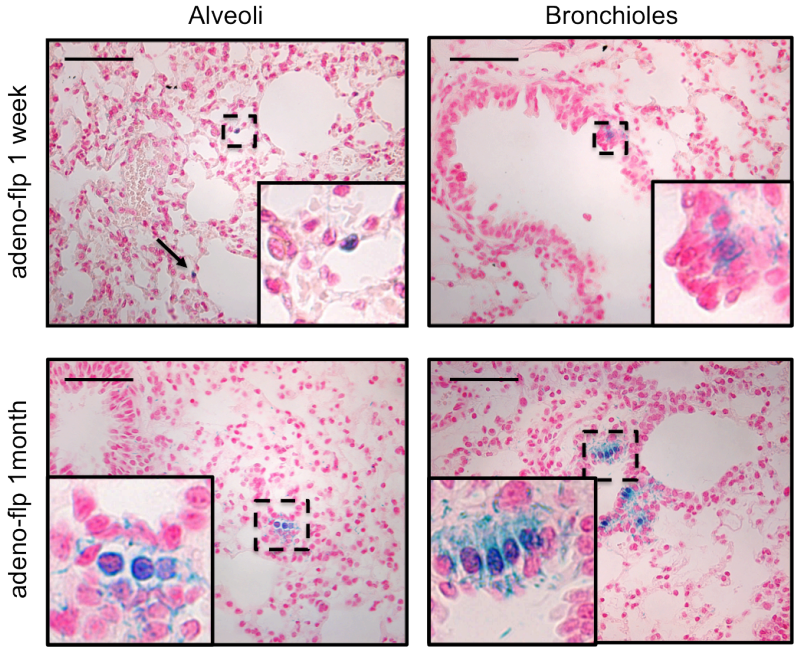


Figure 39 Etv4 is *de novo* expressed in alveolar and bronchiolar cells at early time points following K-Ras^{G12V} activation. Representative pictures from whole mount X-Gal stained lungs of K-Ras$^{+/FSFG12V}$ Etv4$^{+/NLacZ}$ mice one week (**upper panel**) or one month (**lower panel**) after adeno-FlpO infection. Counterstaining with NFR. Scale bars represent 100 µm. Inserts: higher magnification

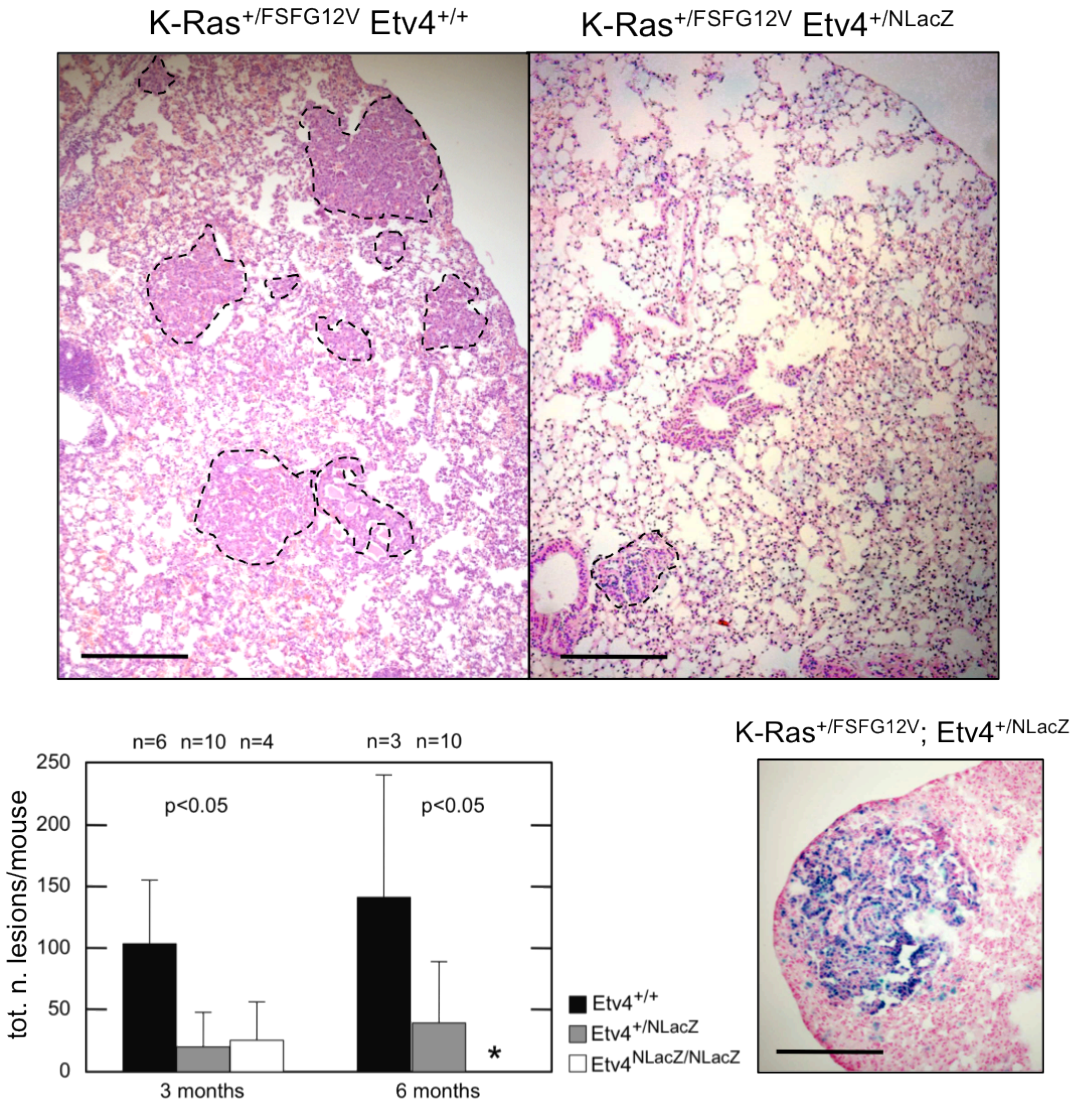


Figure 40: Etv4 is important for K-Ras^{G12V}-driven lung tumorigenesis. Upper panel, representative pictures of lungs from K-Ras$^{+/FSFG12V}$ Etv4$^{+/+}$ (dashed lines indicate hyperplastic or adenomatous regions) and K-Ras$^{+/FSFG12V}$ Etv4$^{+/NlacZ}$ mice (right) 3 months after adeno-FlpO infection. H&E staining. Scale bars represent 200 µm. **Lower left panel,** quantification of lesions (hyperplasias and adenomas). *samples from K-Ras$^{+/FSFG12V}$; Etv4$^{NLacZ/NLacZ}$ mice 6 months after infection are not available yet. **Lower right panel,** one of the few adenomas in a K-Ras$^{+/FSFG12V}$ Etv4$^{+/NlacZ}$ mice. Criosection stained for X-Gal reaction, counterstained with NFR. Scale bar represents 200 µm

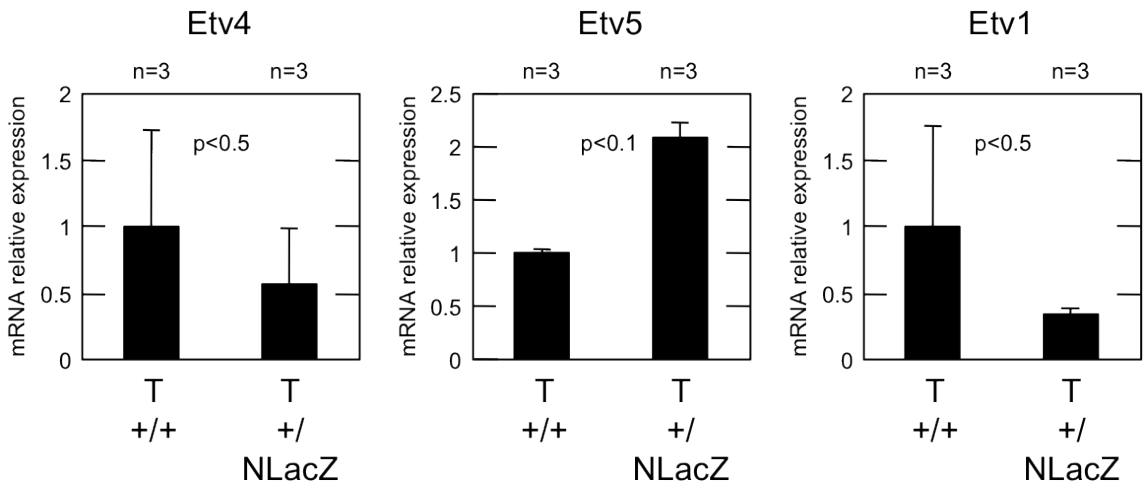
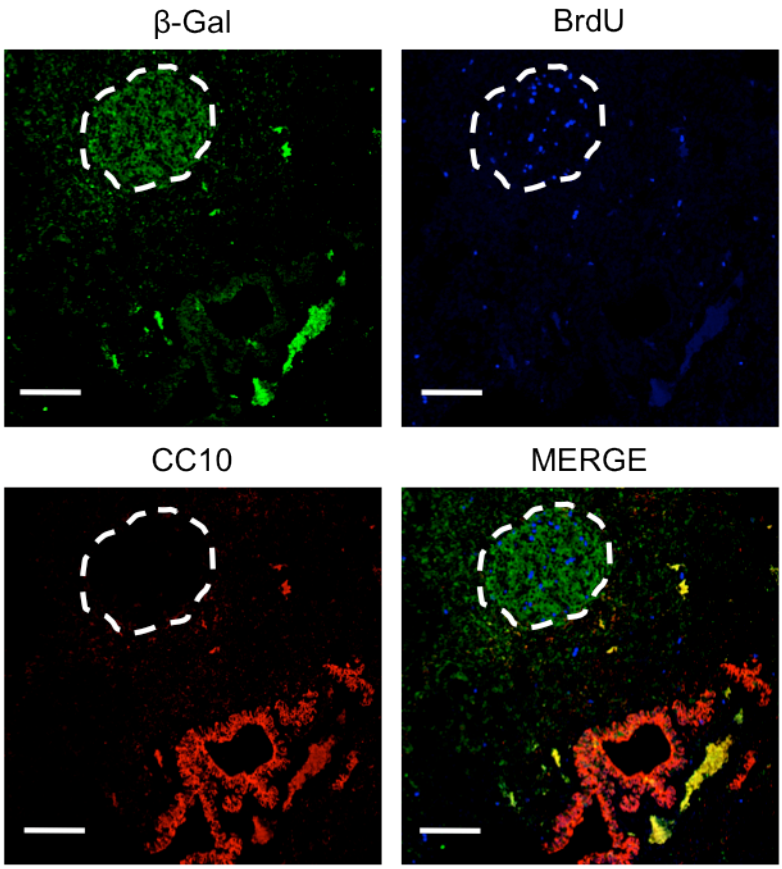
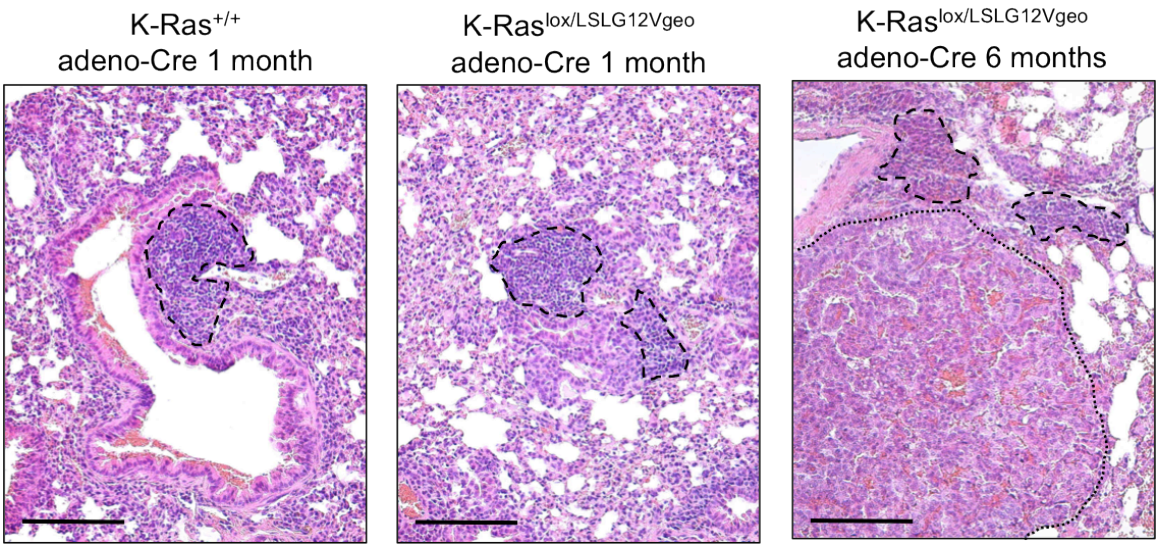


Figure 41: expression of Etv1, Etv4 and Etv5 in K-Ras$^{+/FSFG12V}$ Etv4$^{+/+}$ and K-Ras$^{+/FSFG12V}$ Etv4$^{+/NlacZ}$ tumors. Tumors were isolated from mice of both genotypes 6 months after adeno-FlpO infection. n= number of mice.

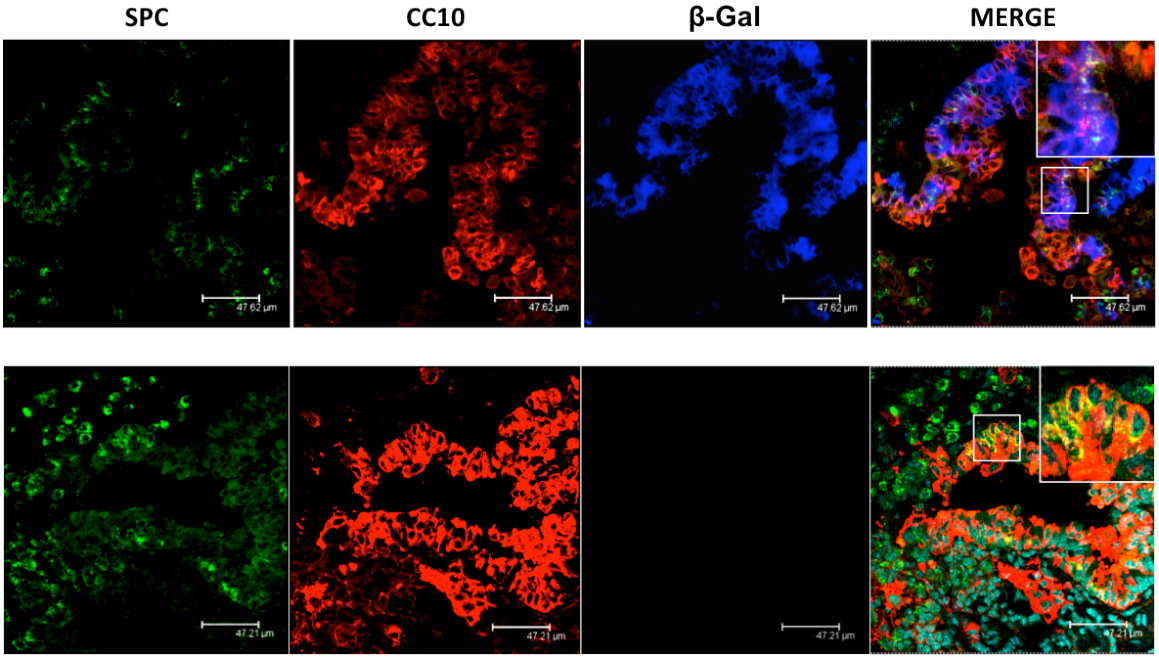
Supplemental Material



Supplemental Figure 1: proliferating cells in K-Ras^{G12V}-induced adenoma. Representative immunofluorescence with antibodies against the β-Galactosidase, BrdU and CC10, performed on lung sections from K-Ras$^{lox/LSLG12Vgeo}$; RERT$^{ert/ert}$ mice treated with a single 4OHT injection and sacrificed 12 weeks later. The mice received BrdU (100 µl/10 gr) every 24 hs for 3 days before they were sacrificed. The dashed line marks the adenoma. Scale bars represent 100 µm.



Supplemental Figure 2: Inflammation in adeno-Cre infected mouse lungs. Representative pictures from HE stained lung sections from mice infected intratracheally with adeno-Cre virus. The genotypes and time points of the sacrifice are indicated. Dashed lines indicate inflammatory cells infiltrates. The dotted line indicates tumor edge. Scale bars represent 200 µm.



Supplemental Figure 3: BASCs expansion 1 month after adeno-Cre intratracheal infection in both K-Ras-positive (upper panel) and K-Ras negative (lower panel) BADJs. Sections from X-Gal whole mount stained lungs of K-Ras$^{lox/LSLG12Vgeo}$ mice sacrificed 1 month after adeno-Cre inoculation. IF using anti-SPC and anti-CC10 antibodies. β-Gal signal acquired in bright field and overlayed. Inserts: higher magnification. Scale bars represent 47 µm.

7 DISCUSSION

Lung cancer, together with cardiovascular diseases, is one of the leading causes of cancer-related deaths in industrialized countries. Over the past 10–15 years, it has become increasingly clear that there is considerable molecular heterogeneity among lung adenocarcinomas. This may be due to the different combinations of genes mutated in these tumors. In fact currently, 26 different genes have been implicated in lung adenocarcinomas (Ding et al., 2008), with 30% of tumors carrying activating K-RAS mutations (Travis et al., 2004). Among tumors with K-RAS activation, another potential source of heterogeneity, with clinical implications for targeted therapy, is the cell type in which the mutation arises. For example, similar K-RAS mutations have been identified in pancreatic cancer, colon cancer, lung cancer, thyroid cancer, and myeloid leukemia (Bos, 1989). However, these tumors have markedly different histologies and behaviors. Such cell of origin effects have not been exhaustively examined for lung cancer, and the cell(s) of origin of human lung adenocarcinoma have not been clearly characterized. In particular, a gene expression profiling of early transformed lung cancer cells has not been reported, while it could provide important information about the molecular events accompanying the first steps of lung tumorigenesis, as well as new markers for lung cancer initiating cells.

The identification of the cellular origins of cancer is crucial to enhance our understanding of the mechanisms regulating the different steps of tumor initiation and progression. Using inducible tumor mouse models, it is now possible to isolate tumor-initiating cells at different time points after the initial oncogenic event, to more precisely define the genetic, epigenetic, transcriptional and post-translational modifications associated with each step of tumor progression. Lineage-tracing experiments have also been used to mark individual tumor cells and assess their respective contribution to tumor growth and relapse after therapy. Such studies demonstrated for example that benign skin and intestine tumors contain cancer stem cells that fuel tumor growth (Driessens et al., 2012); (Schepers et al., 2012) and that the brain tumor glioblastoma contains a discrete population of cells responsible for relapse after therapy (Chen et al., 2012). As for lung cancer, a direct equivalence between cancer initiating cells and cancer

stem cells has not been proven. For this, it is important to identify reliable markers for the cells of origin of lung cancer in order to be able to study their contribution to lung tumor propagation and resistance to therapy.

In the present study, we use unbiased expression of the Ras^{G12V} oncogene achieved by the inducible and ubiquitous expression of the Cre recombinase in the $K\text{-}Ras^{lox/LSLG12Vgeo}$; $RERT^{ert/ert}$ mouse model. Thanks to the β-Galactosidase marker, which is expressed in a bicistronic way together with the K-Ras oncogene, we are able to follow the fate of single $K\text{-}Ras^{G12V}$-expressing cells and to identify the ones capable to start uncontrolled proliferation. Moreover, by complementarily using an adenovirus-mediated approach to induce the Cre recombinase in the lung, we can compare the different effects of a random (RERT) *versus* a bronchiole-directed (adeno-Cre) oncogene activation. Besides, by embryonically inducing the $K\text{-}Ras^{G12V}$ oncoprotein in the $K\text{-}Ras^{+/LSLG12Vgeo}$; $Tg^{Sca1\text{-}Ce}$ model, we are able to study how developmental times determine permissiveness to $K\text{-}Ras^{G12V}$-induced transformation. Last, we used $K\text{-}Ras^{+/LSLG12Vgeo}$; $Rosa26^{LSLEYFP/LSLEYFP}$ mice to isolate early oncogene-responding cells, and to study the changes that occur at the transcriptional level following $K\text{-}Ras^{G12V}$ activation. This will allow us to identify the most relevant genes expressed by the transformed cells, which may be crucial for lung tumor initiation.

Alveolar Type II is the most likely cell type to initiate lung adenocarcinoma upon unbiased $K\text{-}Ras^{G12V}$ expression.

Human K-RAS mutant tumors are almost exclusively adenocarcinomas (Brose et al., 2002); (Suzuki et al., 1990), and they tend to occur peripherally, in the alveolar area, differently from small-cell and squamous tumors that preferentially occur centrally (Giangreco et al., 2007). Moreover, K-RAS mutations were found exclusively in alveolar tumors but not bronchial tumors in a PCR-based analysis of resected human adenocarcinomas (Cooper et al., 1997). Spontaneous lung adenocarcinomas in mice also show a preferential distal localization and morphological and ultrastructural characteristics of ATII cells (Kauffman and Sato, 1985). Moreover, in a chemically

induced mouse model of lung adenocarcinoma, codon 12 K-Ras mutants were identified, and electron microscopy revealed ATII features in the tumor cells (Belinsky et al., 1992). Also with more refined knock in mouse models as the K-Ras^{LSLG12D} (Johnson et al., 2001) or our K-Ras^{LSLG12Vgeo} the tumors obtained widely stain for the ATII marker SPC. Nevertheless, the origin of lung adenocarcinoma from ATII cells cannot be directly inferred from these evidences. In fact, we cannot exclude other possibilities, also considered by other investigator in the field. For instances, oncogene activation could induce hyperproliferation of other cell types, such as Clara cells, which then undergo transdifferentiation as they move into the alveolar compartment. As an alternative model, K-Ras activation could take place in a stem cell with the potential to develop into alveolar type II cells and form hyperplasia and adenomas of such a cell type. Evidences from other cancer types also suggest that tumor differentiation is not necessarily indicative of its cellular origin. Some tumors, such as SCLCs, are indeed derived from neuroendocrine cells as expected from their differentiation (Sutherland et al., 2011), whereas others, such as skin basal cell carcinoma or basal- like breast cancers, originate from cells that normally do not express the markers observed in the resulting tumors (Youssef et al., 2010); (Molyneux et al., 2010).

During the last few years, the issue of the cancer initiating cell for NSCLC has been addressed by several investigators. One approach to the problem has been to exploit some of the inducible mouse models of human lung cancer that have been established (see section 3.3.1). The earliest studies used inhaled adeno-Cre to induce expression of oncogenic K-Ras throughout the respiratory epithelium (Kim et al., 2005). In this model, expansion of double SPC-CC10 positive cells (putative BASCs) at the BADJs was taken as evidence for them being the cells of origin of adenocarcinoma. The expansion of these cells following naphthalene-induced injury, combined with the increased number of tumors after naphthalene treatment, further supported this idea. However, in another study, Xu and colleagues used two different inducible Cre recombinases targeting ATII and the majority of BASCs (Sfpc-CreERT) or Clara cells and some ATII cells (CC10-CreERT) to activate K-Ras^{G12D} and inactivate p53 at the same time. They found that BASCs do

indeed proliferate in response to the two oncogenic events and form hyperplasias, but they do not progress into malignant adenocarcinomas (Xu et al., 2012). In contrast, using Sfpc-CreERT in the ATII cells leads to adenocarcinoma formation, whereas no lesions were found to originate from the BASCs despite the high frequency of genetic recombination in these cells. Overall, this study points at the ATII cells as the cancer initiating cells for lung adenocarcinoma, while Clara cells and BASCs would only be able to give hyperplasias. However, we have to notice that these results were obtained in the absence of p53, an event that generally occurs later during NSCLC progression, and that invariably leads to genetic instability and accumulation of secondary mutations in the tumor cells. We also know from other studies that important features of the tumors, like the tumor-sustaining cell population, are highly dependent on the tumor genotype (Curtis et al., 2010). Thus, we cannot exclude that concomitant p53 deletion may influence the nature of the cells that are transformed upon K-Ras oncogene activation.

In the present work we used the K-Ras^{LSLG12Vgeo} model (Guerra et al., 2003) in which oncogenic K-Ras is the only driving event of tumorigenesis. Of note, our system has the advantage of carrying the bicistronic β-Gal reporter that allows us to follow the individual fate of the K-Ras^{G12V}-expressing cells during the first rounds of cell division, before they yield histologically identifiable lesions such as hyperplasias or adenomas. In this way we can bypass the need of targeting specific cell types, with all the problems concerning the choice of the appropriate markers and promoters, while we can randomly activate the oncogene in all the different cell types and subsequently trace their eventual expansion. This allowed us to compare in the same mice the response of different lung cell types to the K-Ras oncogene expression. By following this approach, we found that all the major cell types (ATII, Clara cells and BASCs) have the capacity to initiate a proliferative response upon expression of an endogenous K-Ras^{G12V} oncoprotein. This may explain the expansion in the number of BASCs observed by Kim et al. when they expressed oncogenic K-Ras^{G12D} (Kim et al., 2005). Nevertheless, thanks to the use of the β-Gal as a surrogate marker of K-Ras^{G12V} expression, we were able to demonstrate for the first time that only ATII cells are able to maintain a sustained

proliferative response, giving rise to big clusters 4 weeks after induction of the oncogene, which will evolve into adenomas and eventually adenocarcinomas. On the contrary Clara cells and BASCs never expanded beyond 20 cells in 2D sections. The reasons for the susceptibility of ATII cells for mutant K-Ras-induced cancer are presently unclear, but following our results we can hypothesize that it relies at least partly in their capacity to upregulate the Ets transcription factor Etv4.

Of note, the genotype that we used in the present study is the K-Ras$^{lox/LSLG12Vgeo}$ in which expression of the oncogenic G12V form of Ras occurs concomitantly with inactivation of the WT allele, mimicking the frequent phenomenon of loss of heterozigosity, although at the initial steps of tumor development (Puyol et al., 2010). We already observed that those mice develop exactly the same number and grade of lesions as the non-LOH K-Ras$^{+/LSLG12Vgeo}$ mice, with the only difference residing in an accelerated tumor development. Thus we decided to use the LOH model in the described work in order to have a faster tool to investigate our hypothesis. However, given recent evidences that lots of the tumor characteristics, including the subpopulation of cells able to sustain tumor growth, depend on the tumor genotype (Curtis et al., 2010), we cannot completely rule out that part of the results we obtained may be specific for the K-Ras$^{lox/LSLG12Vgeo}$ LOH model. Despite this, we strongly think that our data were obtained in the most physiological conditions, closer to the human situation, in which a mutation can occur randomly in all the cells in a given organ, providing a definitive support to previous studies pointing at ATII as the most likely cell of origin of lung adenocarcinoma.

Clara cells can be induced to hyperproliferate following K-Ras^{G12V} expression either by adeno-Cre infection or by expression of the oncogene during embryonic development.

Previous studies by other groups reported that Clara cells are also able to expand following oncogenic K-Ras expression induced by adeno-Cre infection, and that they are able to yield hyperplasias (Xu et al., 2012) but not adenocarcinomas. Yet, when we randomly induced K-Ras^{G12V} expression in K-Ras$^{lox/LSLG12Vgeo}$; RERT$^{ert/ert}$ mice by 4OHT

administration, we could not observe any relevant bronchiolar hyperplasia. Hence, we decided to use in parallel adeno-Cre intratracheal infection in our K-Ras$^{lox/LSLG12Vgeo}$ mice. Surprisingly, we could observe extensive bronchiolar hyperplasia 6 months after infection, in line with the previous reports, as well as adenomas expressing both SPC and the Clara cell marker CC10. In particular, three different cell populations were present at the same time in the same lesion (either hyperplasia or adenoma): CC10-positive cells, SPC-positive cells and CC10-SPC-duble positive cells. The elevated number of this kind of lesions, and their diffuse presence all along the bronchiolar arborizations, and not only at the BADJs where the BASCs reside, makes it unlikely that they arise from multipotent BASCs. We rather believe that transformation of Clara cells may endow them with the capacity to transdifferentiate while expanding. In general, we believe that the capacity of bronchiolar cells to respond to the oncogenic stimulus is also influenced by the system used to activate the oncogene. In other words, we hypothesize that adenoviral infection, that preferentially induces K-Ras^{G12V} expression in the bronchioles, is indeed able to provoke sustained proliferation of Clara cells that, otherwise, would arrest after few replications. In particular, the presence of massive infiltration of inflammatory cells in adeno-Cre infected lungs (**Supplemental Figure 2**), leads to the hypothesis that inflammation could act as a secondary signal that stimulates Clara cell proliferation, possibly through cytokines production. Also, the BASCs expansion observed by Kim et al. following oncogenic K-Ras^{G12D} induction (Kim et al., 2005) may be due to inflammation caused by adeno-Cre infection, as no marker for K-Ras^{G12D} expression was available for them to confirm that BASCs amplification was a direct consequence of the oncogene expression. Indeed, we have evidences that clusters of BASCs that slightly expanded 4 weeks after infection in our K-Ras$^{lox/LSLG12Vgeo}$ mice not necessarily express the β-Gal reporter of K-Ras^{G12V} (**Supplemental Figure 3**). The hypothesis that inflammation plays a major role in lung tumor development has been previously suggested both for patients and mice (Yan et al., 2013); (Vendramini-Costa and Carvalho, 2012);(O'Callaghan et al., 2010). Of note, the essential role of inflammation in K-Ras^{G12V}-induced transformation has already been reported for other K-Ras induced tumors, such as Pancreatic Ductal

Adenocarcinoma, where it was demonstrated that adult pancreatic acinar cells can only be transformed by K-Ras^{G12V} in the context of pancreatitis (Guerra et al., 2007). In this perspective, it will be important to further elucidate the link between lung inflammation and tumor susceptibility, in order to possibly develop new strategies to prevent at least some forms of NSCLC.

On the other hand, the massive expression of K-Ras^{G12V} that we obtained in the bronchioles following adeno-Cre infection, that may be due to a physical limit of the intratracheal infection technique (with the virus reaching first the bronchiolar epithelium) may lead to the activation of the oncogene in a particular less abundant transformation-prone subtype of bronchiolar cells that is not easily targeted through 4OHT-induced Cre activation in the RERT model. This hypothetical bronchiolar subpopulation is unlikely to correspond to the BASCs, as the hyperplasias we observed were not restricted to the BADJs regions. Extensive studies at early time points after adeno-Cre infection, similar to the ones we performed following 4OHT administration in K-Ras$^{lox/LSLG12Vgeo}$; RERT$^{ert/ert}$ mice, will help to elucidate this point. Meanwhile, the presence of three different cell types (SPC-positive, CC10-positive and double SPC-CC10-positive) inside the same lesions (either bronchiolar hyperplasias or adenomas) found in adeno-Cre infected mice, supports the hypothesis that they may arise from a cell with multiple differentiating capacities (that may be either intrinsic or a consequence of K-Ras^{G12V}-induced transformation.). In any case, it would be interesting for the clinical implications that this may have, to determine, for example, whether smoking is more likely to produce mutations in the bronchiolar rather than alveolar epithelium, in a way similar to adeno-Cre intratracheal infection, being the bronchioles the first epithelial barrier to get in contact with the smoking products. It is also known that smoking produces inflammatory changes in the airways of mice, which are only partially reversed after smoking cessation (Braber et al., 2010). More importantly, tobacco smoke promotes tumor development both in carcinogen-treated mice and in transgenic mice undergoing sporadic K-Ras activation in lung epithelial cells (Takahashi et al., 2010).

We also tested whether the capacity of the different lung cell types to be transformed by K-Ras^{G12V} could be determined by the developmental time at which the mutation is switched on, thanks to the Tg$^{Sca1\text{-}Ce}$; K-Ras$^{+/LSLG12Vgeo}$ transgenic model in which the K-Ras oncogene is expressed in the developing lung, as well as in many other organs, starting from E 11.5. At this time point the lung is composed of progenitor cells that express the SPC marker (Liu et al., 2003a); (Warburton et al., 2000). These progenitors will later differentiate proximally into bronchiolar cells and distally into mature type I and type II pneumocytes. At E 18.5 the CC10 marker starts to be expressed in the developing Clara cells and at this time point the Tg$^{Sca1\text{-}Ce}$; K-Ras$^{+/LSLG12Vgeo}$ mice show K-Ras^{G12V} expression in these early bronchiolar cells. Tg$^{Sca1\text{-}Ce}$; K-Ras$^{+/LSLG12Vgeo}$ animals show multiple defects in various organs due to K-Ras^{G12V} expression during embryonic development, and they eventually die around 4 weeks of age. Between their various defects, these mice show massive bronchiolar hyperplasia in the lung, with Clara cells proliferating inside the lumen of the bronchioles inducing obstruction of the airways, causing serious respiratory difficulties in these animals. Thus, Clara cells have clearly undergone massive uncontrolled proliferation in their lungs. We hypothesize that embryonic Clara cells may retain some characteristics that will be lost during adult age, as may be a certain pluripotency, as well as the sensitivity to K-Ras^{G12V}-induced transformation. Remarkably, the notion that susceptibility to K-Ras^{G12V}-driven transformation may depend on the developmental timing was already reported for other K-Ras-induced tumors. In particular it was demonstrated that pancreatic acinar cells can be induced to transform by K-Ras^{G12V}-expression in the embryo but not in the adult mouse, unless concomitant pancreatitis is induced (Guerra et al., 2007)

Early K-Ras^{G12V}-transformed lung cells activate a specific transcriptional program.

In the present work we were able to obtain for the first time the gene expression profile of the cells expanding following K-Ras^{G12V} activation at a very early time point (4 weeks after adeno-Cre infection), at a stage when the transformed cells have only formed small

clusters that are still morphologically indistinguishable from normal lung epithelium. This was possible thanks to the use of the K-Ras$^{lox/LSLG12Vgeo}$; Rosa26$^{LSLEYFP/LSLEYFP}$ mice, in which the EYFP gene acts as a reporter of the Cre recombinase activity. Following confirmation of expansion of the YFP-positive population after K-Ras^{G12V} activation, we could use the YFP fluorescent reporter to sort the expanded cell population. Transcriptome analysis revealed that as early as 4 weeks after the oncogene activation, the targeted cells have already activated a complex and specific transcriptional program. As a control, we observed in the sorted cells a significant upregulation of the Erk/MAPK pathway by gene set enrichment analysis (GSEA, not shown), with important Ras effectors as Grb2, SOS1 and Elk, being significantly upregulated as compared to K-Ras$^{+/+}$ cells. Moreover, pathways associated with cell cycle progression and survival were also enriched, while cell death and apoptosis-related pathways were significantly downregulated. Overall, we found more than 900 significant differentially expressed genes. Of note, the most upregulated gene in K-Ras^{G12V} positive sorted cells was the proto-oncogene Myc. It is already known from other reports, that Myc is strongly oncogenic when overexpressed. Interestingly, using a modified Myc construct, OmoMyc, which acts as a dominant-negative form of Myc and directs the Myc interactome to repression rather than activation (Savino et al., 2011) the group of Gerard Evans was able to show that mutant K-Ras induced lung adenocarcinomas could be eradicated and that the effectiveness could be enhanced by appropriate dosing schedules (Soucek et al., 2008).

The second most relevant transcriptional change in the profile of early K-Ras^{G12V}-transformed cells is the upregulation of the transcription factor Etv4, that we selected for further validation given its previously reported role in lung development as well as in lung cancer progression. But besides this, we believe that the gene expression signature of early K-Ras^{G12V}-transformed cells that we report here can provide many other interesting markers for the study of NSCLC initiating cells, as well as possible new targets for NSCLC treatment.

Etv4 is important, although not essential, for K-Ras^{G12V}-induced lung tumor development.

Etv4 is a member of the Ets family of transcription factors, characterized by the presence of the ETS domain, which is composed of ~ 85 amino acids and binds to DNA sequences with a 5′-GGA(A/T)-3′ core (Sharrocks, 2001). The human ETS family is composed of 28 members, which are differentially expressed in the distinct tissues and cell types. In particular the PEA3 subgroup consists of the highly related Etv1, 4 and 5 proteins. The first member of this subfamily to be cloned was PEA3 (polyomavirus enhancer activator 3), also called E1AF for adenovirus E1A enhancer-binding protein (Xin et al., 1992); (Higashino et al., 1993). The rational name for this protein is nowadays Etv4 (ETS variant 4). Two more relatives were identified soon thereafter, Etv1 (also called ER81 for ETS- related 81) and Etv5 (also called ERM for ETS-related molecule) (Brown and McKnight, 1992); (Monte et al., 1994). The three proteins are expressed in numerous organs both during embryonic development and in adults (Hollenhorst et al., 2004); (Chotteau-Lelievre et al., 1997); (Chotteau-Lelievre et al., 2001). One emerging role of PEA3 factors is in branching morphogenesis, where primitive epithelial buds bifurcate to generate arborized ducts or acinar structures. In particular, in the developing mouse lung, expression of Etv4 and 5 is initially restricted to the distal buds: Etv5 is transcribed exclusively in the epithelium, while Etv4 is expressed in both epithelium and mesenchyme (Liu et al., 2003b). All Pea3 sub-family members have been genetically inactivated in mice. Etv5 null mutant dies at E9.5, before lung development starts (Liu et al., 2003b). Etv4 null male mice are infertile, probably because of neuronal dysfunction, but apparently do not have a lung phenotype (Laing et al., 2000), suggesting that Etv4 and Etv5 have redundant functions during lung development. Besides their role in normal embryonic development, Pea3 family members have also been implicated in cancer. In Ewing tumors, chromosomal translocations affecting Etv1 or Etv4 are an underlying cause of carcinogenesis (Janknecht, 2005). Likewise, chromosomal rearrangements of the Etv1, Etv4 or Etv5 gene occur in prostate tumors and are thought to be one of the major driving forces in the genesis of prostate cancer

(Tomlins et al., 2005); (Tomlins et al., 2006); (Helgeson et al., 2008). In addition, these three ETS proteins are implicated in melanomas, breast and other types of cancer (Jane-Valbuena et al., 2010); (Trimble et al., 1993); (Shepherd et al., 2001). In particular in NSCLC Etv4 has been found to be overexpressed at the mRNA level and appears to promote cell motility and invasion (Hiroumi et al., 2001). However, Etv4 was never associated with tumor initiation before. Adding an important piece of information, our data suggest that Etv4 upregulation is one of the first and most significant transcriptional events accompanying the very early steps of K-Ras^{G12V}-induced transformation. Etv4 overexpression is maintained during lung tumor progression, possibly conferring a growth advantage to the clones that express it. Interestingly, elimination of a single Etv4 allele significantly reduces the number of initial lesions 3 months after induction of the oncogene, underlying the important role of this transcription factor as a downstream mediator of K-Ras^{G12V}-induced transformation. Interestingly, no big differences could be observed between the number of pretumoral lesions in Etv4 heterozygous and KO mice, proving that Etv4 is not completely essential for K-Ras^{G12V}-driven tumorigenesis. Interestingly, tumors isolated from K-Ras$^{+/FSFG12V}$; Etv4$^{+/NLacZ}$ mice 6 months after adeno-FlpO infection show increased expression of the other member of the Pea3 subgroup of Ets factors Etv5, leading to the hypothesis of a possible compensatory effect between these highly related proteins. In this perspective, it will be interesting to determine if a compensatory mechanism, possibly involving upregulation of Etv5, are activated by the tumor cells to overcome the complete lack of Etv4 in the K-Ras$^{+/FSFG12V}$; Etv4$^{NLacZ/NLacZ}$ mice. In any case, our data point at Etv4 as a key mediator of tumorigenic K-Ras downstream signaling. Previous reports postulated a possible link between K-Ras and Etv4 by demonstrating that Etv4 expression is controlled by the transcriptional repressor Capicua in melanoma cells, and that Erk phosphorilation following MAPK pathway activation leads to Capicua's phosphorilation and release from the Etv4 promoter (Dissanayake et al., 2011). Thus it would be interesting to assess whether a similar mechanism is also responsible for Etv4 upregulation during the early steps of K-Ras^{G12V}-driven transformation in lung.

Moreover, it will be crucial to determine the genes whose expression is controlled by Etv4 once it is upregulated following MAPK pathway activation. In fact, so far, the most abundantly described targets of Etv4 have been the metalloproteases family genes (Higashino et al., 1995); (Shindoh et al., 1996), (Hiroumi et al., 2001), which are responsible for degradation of the extracellular matrix, and for this reason Etv4 has always been associated with cancer invasion and metastasis, but not tumor initiation. We believe that other important Etv4 target genes that play a role in tumor initiation (possibly related to proliferation and survival) still need to be discovered and may hopefully provide interesting targets for new therapeutic strategies. In this perspectives, ChIP-seq experiments on Etv4 overexpressing lung cancer cells are needed and will hopefully be highly informative. In particular, Etv4 as any other transcription factor is not considered as a good druggable therapeutic target, but it may control downstream pathways that offer indeed a possible therapeutic strategy. A similar approach (integration of gene expression and gene occupancy analysis) already led to the discovery that combined inhibition of proteasome and Rho/Rock pathways, both controlled by the transcription factor Gata2, causes marked tumor clearance in a K-Ras driven model of NSCLC (Kumar et al., 2012).

Overall, our work identifies a major effector downstream of the MAPK pathway which is rapidly activated following expression of the oncogenic form of K-Ras, and which is probably driving the transcription of a set of clue genes whose identity will be essential to discover.

8 CONCLUSIONS

THE CONCLUSIONS OBTAINED IN THIS WORK WERE THE FOLLOWING:

1. In the mouse lung, Alevolar Type II cells are the most likely cell type to initiate lung adenocarcinoma following random, unbiased K-Ras^{G12V} activation.
2. Clara cells and BASCs also respond to K-Ras^{G12V} activation with an initial proliferative response, but they are unable to give hyperplasias or adenomas.
3. Clara cells can be induced to give hyperplasias and eventually adenomas when K-Ras^{G12V} is preferentially induced in the bronchiolar epithelium following adeno-Cre intratracheal infection and the consequent inflammatory response.
4. Clara cells can be induced to give hyperplasias and eventually adenomas when K-Ras^{G12V} is expressed early during lung embryonic development.
5. The cells that start to expand following K-Ras^{G12V} activation quickly activate a specific transcriptional program, characterized by the upregulation of the Ets transcription factor Etv4.
6. Etv4 is *de novo* expressed in bronchioles and alveoli following K-Ras^{G12V} activation.
7. Elimination of one Etv4 allele reduces the number of tumors induced by K-Ras^{G12V}.
8. Elimination of both Etv4 alleles does not completely impede tumor formation, possibly due to Etv5 compensatory upregultaion.

LAS CONCLUSIONES OBTENIDAS EN ESTE TRABAJO FUERON:

1. En el pulmón de ratón, las células alveolares de tipo II son las que con más probabilidad pueden iniciar los tumores inducidos por K-Ras^{G12V}, puesto que la activación del oncogen se realice de forma aleatoria y no sea preferentemente dirigida a ninguno de los epitelios que forman el pulmón.
2. Las células Clara y las células madres bronquioalveolares (BASCs) también responden a la activación del oncogen K-Ras con una initial proliferación, sin embargo son incapables de dar lugar a hiperplasias o adenomas.
3. Las células Clara pueden dar lugar a hiperplasias y eventualmente adenomas en el caso que se fuerze la expresión del oncogen K-Ras en el epitelio bronquiolar a través de la infección intratraqueal con el adenovirus adeno-Cre, la cual conlleva una respuesta inflamatoria en el pulmón.
4. Las células Clara pueden dar lugar a hiperplasias y eventualmente adenomas en el caso que la expresión del oncogen K-Ras se produzca durante el desarrollo embrionario del pulmón.
5. La población celular que se expande tras la expresión del oncogen K-Ras activa rapidamente un programa de trasncripción genica especifico, que se caracteriza por el aumento de expresión del facor de trasncripción de la familia Ets, Etv4.
6. Etv4 se expresa *de novo* en los epitelios bronquiolar y alveolar tras la activación del oncogen K-Ras.
7. La eliminación de un alelo de Etv4 disminuye el número de tumores que se producen tras la activación de K-Ras^{G12V}.
8. La completa eliminación de Etv4 no es capaz de prevenir de forma completa el desarrollo te tumores inducidos por K-Ras^{G12V}, probablemente a causa de una compensación debida a la sobrexpresión del factor relacionado Etv5.

9 BIBLIOGRAPHY

Al-Hajj, M., Wicha, M. S., Benito-Hernandez, A., Morrison, S. J., and Clarke, M. F. (2003). Prospective identification of tumorigenic breast cancer cells. Proc Natl Acad Sci U S A *100*, 3983-3988.

Ballester, R., Marchuk, D., Boguski, M., Saulino, A., Letcher, R., Wigler, M., and Collins, F. (1990). The NF1 locus encodes a protein functionally related to mammalian GAP and yeast IRA proteins. Cell *63*, 851-859.

Beasley, M. B., Lantuejoul, S., Abbondanzo, S., Chu, W. S., Hasleton, P. S., Travis, W. D., and Brambilla, E. (2003). The P16/cyclin D1/Rb pathway in neuroendocrine tumors of the lung. Hum Pathol *34*, 136-142.

Belinsky, S. A., Devereux, T. R., Foley, J. F., Maronpot, R. R., and Anderson, M. W. (1992). Role of the alveolar type II cell in the development and progression of pulmonary tumors induced by 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone in the A/J mouse. Cancer Res *52*, 3164-3173.

Blanpain, C. (2013). Tracing the cellular origin of cancer. Nat Cell Biol *15*, 126-134.

Bos, J. L. (1989). ras oncogenes in human cancer: a review. Cancer Res *49*, 4682-4689.

Braber, S., Henricks, P. A., Nijkamp, F. P., Kraneveld, A. D., and Folkerts, G. (2010). Inflammatory changes in the airways of mice caused by cigarette smoke exposure are only partially reversed after smoking cessation. Respir Res *11*, 99.

Brocard, J., Warot, X., Wendling, O., Messaddeq, N., Vonesch, J. L., Chambon, P., and Metzger, D. (1997). Spatio-temporally controlled site-specific somatic mutagenesis in the mouse. Proc Natl Acad Sci U S A *94*, 14559-14563.

Broek, D., Toda, T., Michaeli, T., Levin, L., Birchmeier, C., Zoller, M., Powers, S., and Wigler, M. (1987). The S. cerevisiae CDC25 gene product regulates the RAS/adenylate cyclase pathway. Cell *48*, 789-799.

Brose, M. S., Volpe, P., Feldman, M., Kumar, M., Rishi, I., Gerrero, R., Einhorn, E., Herlyn, M., Minna, J., Nicholson, A*., et al.* (2002). BRAF and RAS mutations in human lung cancer and melanoma. Cancer Res *62*, 6997-7000.

Brown, T. A., and McKnight, S. L. (1992). Specificities of protein-protein and protein-DNA interaction of GABP alpha and two newly defined ets-related proteins. Genes Dev *6*, 2502-2512.

Capon, D. J., Seeburg, P. H., McGrath, J. P., Hayflick, J. S., Edman, U., Levinson, A. D., and Goeddel, D. V. (1983). Activation of Ki-ras2 gene in human colon and lung carcinomas by two different point mutations. Nature *304*, 507-513.

Chen, J., Li, Y., Yu, T. S., McKay, R. M., Burns, D. K., Kernie, S. G., and Parada, L. F. (2012). A restricted cell population propagates glioblastoma growth after chemotherapy. Nature *488*, 522-526.

Chotteau-Lelievre, A., Desbiens, X., Pelczar, H., Defossez, P. A., and de Launoit, Y. (1997). Differential expression patterns of the PEA3 group transcription factors through murine embryonic development. Oncogene *15*, 937-952.

Chotteau-Lelievre, A., Dolle, P., Peronne, V., Coutte, L., de Launoit, Y., and Desbiens, X. (2001). Expression patterns of the Ets transcription factors from the PEA3 group during early stages of mouse development. Mech Dev *108*, 191-195.

Clarke, M. F., and Fuller, M. (2006). Stem cells and cancer: two faces of eve. Cell *124*, 1111-1115.

Cooper, C. A., Carby, F. A., Bubb, V. J., Lamb, D., Kerr, K. M., and Wyllie, A. H. (1997). The pattern of K-ras mutation in pulmonary adenocarcinoma defines a new pathway of tumour development in the human lung. J Pathol *181*, 401-404.

Curtis, S. J., Sinkevicius, K. W., Li, D., Lau, A. N., Roach, R. R., Zamponi, R., Woolfenden, A. E., Kirsch, D. G., Wong, K. K., and Kim, C. F. (2010). Primary tumor genotype is an important determinant in identification of lung cancer propagating cells. Cell Stem Cell *7*, 127-133.

Dankort, D., Filenova, E., Collado, M., Serrano, M., Jones, K., and McMahon, M. (2007). A new mouse model to explore the initiation, progression, and therapy of BRAFV600E-induced lung tumors. Genes Dev *21*, 379-384.

Ding, L., Getz, G., Wheeler, D. A., Mardis, E. R., McLellan, M. D., Cibulskis, K., Sougnez, C., Greulich, H., Muzny, D. M., Morgan, M. B.*, et al.* (2008). Somatic mutations affect key pathways in lung adenocarcinoma. Nature *455*, 1069-1075.

Dissanayake, K., Toth, R., Blakey, J., Olsson, O., Campbell, D. G., Prescott, A. R., and MacKintosh, C. (2011). ERK/p90(RSK)/14-3-3 signalling has an impact on expression of PEA3 Ets transcription factors via the transcriptional repressor capicua. Biochem J *433*, 515-525.

Dow, L. E., Premsrirut, P. K., Zuber, J., Fellmann, C., McJunkin, K., Miething, C., Park, Y., Dickins, R. A., Hannon, G. J., and Lowe, S. W. (2012). A pipeline for the generation of shRNA transgenic mice. Nat Protoc *7*, 374-393.

Downward, J. (2003). Targeting RAS signalling pathways in cancer therapy. Nat Rev Cancer *3*, 11-22.

Downward, J., Riehl, R., Wu, L., and Weinberg, R. A. (1990). Identification of a nucleotide exchange-promoting activity for p21ras. Proc Natl Acad Sci U S A *87*, 5998-6002.

Driessens, G., Beck, B., Caauwe, A., Simons, B. D., and Blanpain, C. (2012). Defining the mode of tumour growth by clonal analysis. Nature *488*, 527-530.

Essers, M. A., Weijzen, S., de Vries-Smits, A. M., Saarloos, I., de Ruiter, N. D., Bos, J. L., and Burgering, B. M. (2004). FOXO transcription factor activation by oxidative stress mediated by the small GTPase Ral and JNK. EMBO J *23*, 4802-4812.

Fisher, G. H., Wellen, S. L., Klimstra, D., Lenczowski, J. M., Tichelaar, J. W., Lizak, M. J., Whitsett, J. A., Koretsky, A., and Varmus, H. E. (2001). Induction and apoptotic regression of lung adenocarcinomas by regulation of a K-Ras transgene in the presence and absence of tumor suppressor genes. Genes Dev *15*, 3249-3262.

Friedrich, G., and Soriano, P. (1991). Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. Genes Dev *5*, 1513-1523.

George, D. L., Scott, A. F., Trusko, S., Glick, B., Ford, E., and Dorney, D. J. (1985). Structure and expression of amplified cKi-ras gene sequences in Y1 mouse adrenal tumor cells. EMBO J *4*, 1199-1203.

Gerber, D. E., and Minna, J. D. (2010). ALK inhibition for non-small cell lung cancer: from discovery to therapy in record time. Cancer Cell *18*, 548-551.

Giangreco, A., Groot, K. R., and Janes, S. M. (2007). Lung cancer and lung stem cells: strange bedfellows? Am J Respir Crit Care Med *175*, 547-553.

Gibbs, J. B., Schaber, M. D., Allard, W. J., Sigal, I. S., and Scolnick, E. M. (1988). Purification of ras GTPase activating protein from bovine brain. Proc Natl Acad Sci U S A *85*, 5026-5030.

Graziano, S. L., Gamble, G. P., Newman, N. B., Abbott, L. Z., Rooney, M., Mookherjee, S., Lamb, M. L., Kohman, L. J., and Poiesz, B. J. (1999). Prognostic significance of K-ras codon 12 mutations in patients with resected stage I and II non-small-cell lung cancer. J Clin Oncol *17*, 668-675.

Gu, H., Zou, Y. R., and Rajewsky, K. (1993). Independent control of immunoglobulin switch recombination at individual switch regions evidenced through Cre-loxP-mediated gene targeting. Cell *73*, 1155-1164.

Guerra, C., Mijimolle, N., Dhawahir, A., Dubus, P., Barradas, M., Serrano, M., Campuzano, V., and Barbacid, M. (2003). Tumor induction by an endogenous K-ras oncogene is highly dependent on cellular context. Cancer Cell *4*, 111-120.

Guerra, C., Schuhmacher, A. J., Canamero, M., Grippo, P. J., Verdaguer, L., Perez-Gallego, L., Dubus, P., Sandgren, E. P., and Barbacid, M. (2007). Chronic pancreatitis is essential for induction of pancreatic ductal adenocarcinoma by K-Ras oncogenes in adult mice. Cancer Cell *11*, 291-302.

Gupta, P. B., Chaffer, C. L., and Weinberg, R. A. (2009). Cancer stem cells: mirage or reality? Nat Med *15*, 1010-1012.

Hanafusa, H., Torii, S., Yasunaga, T., and Nishida, E. (2002). Sprouty1 and Sprouty2 provide a control mechanism for the Ras/MAPK signalling pathway. Nat Cell Biol *4*, 850-858.

Hancock, J. F. (2003). Ras proteins: different signals from different locations. Nat Rev Mol Cell Biol *4*, 373-384.

Helgeson, B. E., Tomlins, S. A., Shah, N., Laxman, B., Cao, Q., Prensner, J. R., Cao, X., Singla, N., Montie, J. E., Varambally, S*., et al.* (2008). Characterization of TMPRSS2:ETV5 and SLC45A3:ETV5 gene fusions in prostate cancer. Cancer Res *68*, 73-80.

Higashino, F., Yoshida, K., Fujinaga, Y., Kamio, K., and Fujinaga, K. (1993). Isolation of a cDNA encoding the adenovirus E1A enhancer binding protein: a new human member of the ets oncogene family. Nucleic Acids Res *21*, 547-553.

Higashino, F., Yoshida, K., Noumi, T., Seiki, M., and Fujinaga, K. (1995). Ets-related protein E1A-F can activate three different matrix metalloproteinase gene promoters. Oncogene *10*, 1461-1463.

Hiroumi, H., Dosaka-Akita, H., Yoshida, K., Shindoh, M., Ohbuchi, T., Fujinaga, K., and Nishimura, M. (2001). Expression of E1AF/PEA3, an Ets-related transcription factor in human non-small-cell lung cancers: its relevance in cell motility and invasion. Int J Cancer *93*, 786-791.

Hollenhorst, P. C., Jones, D. A., and Graves, B. J. (2004). Expression profiles frame the promoter specificity dilemma of the ETS family of transcription factors. Nucleic Acids Res *32*, 5693-5702.

Horio, Y., Chen, A., Rice, P., Roth, J. A., Malkinson, A. M., and Schrump, D. S. (1996). Ki-ras and p53 mutations are early and late events, respectively, in urethane-induced pulmonary carcinogenesis in A/J mice. Mol Carcinog *17*, 217-223.

Huijbers, I. J., Krimpenfort, P., Berns, A., and Jonkers, J. (2011). Rapid validation of cancer genes in chimeras derived from established genetically engineered mouse models. Bioessays *33*, 701-710.

Jackson, E. L., Willis, N., Mercer, K., Bronson, R. T., Crowley, D., Montoya, R., Jacks, T., and Tuveson, D. A. (2001). Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. Genes Dev *15*, 3243-3248.

Jane-Valbuena, J., Widlund, H. R., Perner, S., Johnson, L. A., Dibner, A. C., Lin, W. M., Baker, A. C., Nazarian, R. M., Vijayendran, K. G., Sellers, W. R*., et al.* (2010). An oncogenic role for ETV1 in melanoma. Cancer Res *70*, 2075-2084.

Janknecht, R. (2005). EWS-ETS oncoproteins: the linchpins of Ewing tumors. Gene *363*, 1-14.

Johnson, L., Mercer, K., Greenbaum, D., Bronson, R. T., Crowley, D., Tuveson, D. A., and Jacks, T. (2001). Somatic activation of the K-ras oncogene causes early onset lung cancer in mice. Nature *410*, 1111-1116.

Jonkers, J., and Berns, A. (2002). Conditional mouse models of sporadic cancer. Nat Rev Cancer *2*, 251-265.

Karnoub, A. E., and Weinberg, R. A. (2008). Ras oncogenes: split personalities. Nat Rev Mol Cell Biol *9*, 517-531.

Kim, C. F., Jackson, E. L., Woolfenden, A. E., Lawrence, S., Babar, I., Vogel, S., Crowley, D., Bronson, R. T., and Jacks, T. (2005). Identification of bronchioalveolar stem cells in normal lung and lung cancer. Cell *121*, 823-835.

Krasilnikov, M. A. (2000). Phosphatidylinositol-3 kinase dependent pathways: the role in control of cell growth, survival, and malignant transformation. Biochemistry (Mosc) *65*, 59-67.

Kuhn, R., Schwenk, F., Aguet, M., and Rajewsky, K. (1995). Inducible gene targeting in mice. Science *269*, 1427-1429.

Kumar, M. S., Hancock, D. C., Molina-Arcas, M., Steckel, M., East, P., Diefenbacher, M., Armenteros-Monterroso, E., Lassailly, F., Matthews, N., Nye, E.*, et al.* (2012). The GATA2 transcriptional network is requisite for RAS oncogene-driven non-small cell lung cancer. Cell *149*, 642-655.

Kwon, M. C., and Berns, A. (2013). Mouse models for lung cancer. Mol Oncol *7*, 165-177.

Laing, M. A., Coonrod, S., Hinton, B. T., Downie, J. W., Tozer, R., Rudnicki, M. A., and Hassell, J. A. (2000). Male sexual dysfunction in mice bearing targeted mutant alleles of the PEA3 ets gene. Mol Cell Biol *20*, 9337-9345.

Lapidot, T., Sirard, C., Vormoor, J., Murdoch, B., Hoang, T., Caceres-Cortes, J., Minden, M., Paterson, B., Caligiuri, M. A., and Dick, J. E. (1994). A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. Nature *367*, 645-648.

Li, J., Zhang, Z., Dai, Z., Plass, C., Morrison, C., Wang, Y., Wiest, J. S., Anderson, M. W., and You, M. (2003). LOH of chromosome 12p correlates with Kras2 mutation in non-small cell lung cancer. Oncogene *22*, 1243-1246.

Liu, C., Ikegami, M., Stahlman, M. T., Dey, C. R., and Whitsett, J. A. (2003a). Inhibition of alveolarization and altered pulmonary mechanics in mice expressing GATA-6. Am J Physiol Lung Cell Mol Physiol *285*, L1246-1254.

Liu, X., Driskell, R. R., and Engelhardt, J. F. (2006). Stem cells in the lung. Methods Enzymol *419*, 285-321.

Liu, Y., Jiang, H., Crawford, H. C., and Hogan, B. L. (2003b). Role for ETS domain transcription factors Pea3/Erm in mouse lung development. Dev Biol *261*, 10-24.

Livet, J., Sigrist, M., Stroebel, S., De Paola, V., Price, S. R., Henderson, C. E., Jessell, T. M., and Arber, S. (2002). ETS gene Pea3 controls the central position and terminal arborization of specific motor neuron pools. Neuron *35*, 877-892.

Lynch, T. J., Bell, D. W., Sordella, R., Gurubhagavatula, S., Okimoto, R. A., Brannigan, B. W., Harris, P. L., Haserlat, S. M., Supko, J. G., Haluska, F. G.*, et al.* (2004). Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med *350*, 2129-2139.

Malkinson, A. M. (1989). The genetic basis of susceptibility to lung tumors in mice. Toxicology *54*, 241-271.

Malkinson, A. M. (2001). Primary lung tumors in mice as an aid for understanding, preventing, and treating human adenocarcinoma of the lung. Lung Cancer *32*, 265-279.

Malumbres, M., and Barbacid, M. (2003). RAS oncogenes: the first 30 years. Nat Rev Cancer *3*, 459-465.

Mao, X., Fujiwara, Y., and Orkin, S. H. (1999). Improved reporter strain for monitoring Cre recombinase-mediated DNA excisions in mice. Proc Natl Acad Sci U S A *96*, 5037-5042.

Marshall, C. (1999). How do small GTPase signal transduction pathways regulate cell cycle entry? Curr Opin Cell Biol *11*, 732-736.

Martin, G. A., Viskochil, D., Bollag, G., McCabe, P. C., Crosier, W. J., Haubruck, H., Conroy, L., Clark, R., O'Connell, P., Cawthon, R. M., and et al. (1990). The GAP-related domain of the neurofibromatosis type 1 gene product interacts with ras p21. Cell *63*, 843-849.

McGrath, J. P., Capon, D. J., Smith, D. H., Chen, E. Y., Seeburg, P. H., Goeddel, D. V., and Levinson, A. D. (1983). Structure and organization of the human Ki-ras proto-oncogene and a related processed pseudogene. Nature *304*, 501-506.

Meuwissen, R., and Berns, A. (2005). Mouse models for human lung cancer. Genes Dev *19*, 643-664.

Meuwissen, R., Linn, S. C., Linnoila, R. I., Zevenhoven, J., Mooi, W. J., and Berns, A. (2003). Induction of small cell lung cancer by somatic inactivation of both Trp53 and Rb1 in a conditional mouse model. Cancer Cell *4*, 181-189.

Meuwissen, R., Linn, S. C., van der Valk, M., Mooi, W. J., and Berns, A. (2001). Mouse model for lung tumorigenesis through Cre/lox controlled sporadic activation of the K-Ras oncogene. Oncogene *20*, 6551-6558.

Molloy, C. J., Bottaro, D. P., Fleming, T. P., Marshall, M. S., Gibbs, J. B., and Aaronson, S. A. (1989). PDGF induction of tyrosine phosphorylation of GTPase activating protein. Nature *342*, 711-714.

Molyneux, G., Geyer, F. C., Magnay, F. A., McCarthy, A., Kendrick, H., Natrajan, R., Mackay, A., Grigoriadis, A., Tutt, A., Ashworth, A*., et al.* (2010). BRCA1 basal-like breast cancers originate from luminal epithelial progenitors and not from basal stem cells. Cell Stem Cell *7*, 403-417.

Monte, D., Baert, J. L., Defossez, P. A., de Launoit, Y., and Stehelin, D. (1994). Molecular cloning and characterization of human ERM, a new member of the Ets family closely related to mouse PEA3 and ER81 transcription factors. Oncogene *9*, 1397-1406.

O'Callaghan, D. S., O'Donnell, D., O'Connell, F., and O'Byrne, K. J. (2010). The role of inflammation in the pathogenesis of non-small cell lung cancer. J Thorac Oncol *5*, 2024-2036.

Paez, J. G., Janne, P. A., Lee, J. C., Tracy, S., Greulich, H., Gabriel, S., Herman, P., Kaye, F. J., Lindeman, N., Boggon, T. J*., et al.* (2004). EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science *304*, 1497-1500.

Pao, W., Miller, V., Zakowski, M., Doherty, J., Politi, K., Sarkaria, I., Singh, B., Heelan, R., Rusch, V., Fulton, L*., et al.* (2004a). EGF receptor gene mutations are common in lung cancers from

"never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. Proc Natl Acad Sci U S A *101*, 13306-13311.

Pao, W., Miller, V. A., Venkatraman, E., and Kris, M. G. (2004b). Predicting sensitivity of non-small-cell lung cancer to gefitinib: is there a role for P-Akt? J Natl Cancer Inst *96*, 1117-1119.

Pao-Chun, L., Chan, P. M., Chan, W., and Manser, E. (2009). Cytoplasmic ACK1 interaction with multiple receptor tyrosine kinases is mediated by Grb2: an analysis of ACK1 effects on Axl signaling. J Biol Chem *284*, 34954-34963.

Park, K. S., Liang, M. C., Raiser, D. M., Zamponi, R., Roach, R. R., Curtis, S. J., Walton, Z., Schaffer, B. E., Roake, C. M., Zmoos, A. F.*, et al.* (2011). Characterization of the cell of origin for small cell lung cancer. Cell Cycle *10*, 2806-2815.

Pells, S., Divjak, M., Romanowski, P., Impey, H., Hawkins, N. J., Clarke, A. R., Hooper, M. L., and Williamson, D. J. (1997). Developmentally-regulated expression of murine K-ras isoforms. Oncogene *15*, 1781-1786.

Puyol, M., Martin, A., Dubus, P., Mulero, F., Pizcueta, P., Khan, G., Guerra, C., Santamaria, D., and Barbacid, M. (2010). A synthetic lethal interaction between K-Ras oncogenes and Cdk4 unveils a therapeutic strategy for non-small cell lung carcinoma. Cancer Cell *18*, 63-73.

Rawlins, E. L., Okubo, T., Xue, Y., Brass, D. M., Auten, R. L., Hasegawa, H., Wang, F., and Hogan, B. L. (2009). The role of Scgb1a1+ Clara cells in the long-term maintenance and repair of lung airway, but not alveolar, epithelium. Cell Stem Cell *4*, 525-534.

Re, F. C., Manenti, G., Borrello, M. G., Colombo, M. P., Fisher, J. H., Pierotti, M. A., Della Porta, G., and Dragani, T. A. (1992). Multiple molecular alterations in mouse lung tumors. Mol Carcinog *5*, 155-160.

Regales, L., Balak, M. N., Gong, Y., Politi, K., Sawai, A., Le, C., Koutcher, J. A., Solit, D. B., Rosen, N., Zakowski, M. F., and Pao, W. (2007). Development of new mouse lung tumor models expressing EGFR T790M mutants associated with clinical resistance to kinase inhibitors. PLoS One *2*, e810.

Robinson, L. C., Gibbs, J. B., Marshall, M. S., Sigal, I. S., and Tatchell, K. (1987). CDC25: a component of the RAS-adenylate cyclase pathway in Saccharomyces cerevisiae. Science *235*, 1218-1221.

Rock, J. R., and Hogan, B. L. (2011). Epithelial progenitor cells in lung development, maintenance, repair, and disease. Annu Rev Cell Dev Biol *27*, 493-512.

Rock, J. R., Onaitis, M. W., Rawlins, E. L., Lu, Y., Clark, C. P., Xue, Y., Randell, S. H., and Hogan, B. L. (2009). Basal cells as stem cells of the mouse trachea and human airway epithelium. Proc Natl Acad Sci U S A *106*, 12771-12775.

Savino, M., Annibali, D., Carucci, N., Favuzzi, E., Cole, M. D., Evan, G. I., Soucek, L., and Nasi, S. (2011). The action mechanism of the Myc inhibitor termed Omomyc may give clues on how to target Myc for cancer therapy. PLoS One *6*, e22284.

Schepers, A. G., Snippert, H. J., Stange, D. E., van den Born, M., van Es, J. H., van de Wetering, M., and Clevers, H. (2012). Lineage tracing reveals Lgr5+ stem cell activity in mouse intestinal adenomas. Science *337*, 730-735.

Sharrocks, A. D. (2001). The ETS-domain transcription factor family. Nat Rev Mol Cell Biol *2*, 827-837.

Shepherd, T. G., Kockeritz, L., Szrajber, M. R., Muller, W. J., and Hassell, J. A. (2001). The pea3 subfamily ets genes are required for HER2/Neu-mediated mammary oncogenesis. Curr Biol *11*, 1739-1748.

Shimizu, K., Birnbaum, D., Ruley, M. A., Fasano, O., Suard, Y., Edlund, L., Taparowsky, E., Goldfarb, M., and Wigler, M. (1983). Structure of the Ki-ras gene of the human lung carcinoma cell line Calu-1. Nature *304*, 497-500.

Shimkin, M. B., and Stoner, G. D. (1975). Lung tumors in mice: application to carcinogenesis bioassay. Adv Cancer Res *21*, 1-58.

Shindoh, M., Higashino, F., Kaya, M., Yasuda, M., Funaoka, K., Hanzawa, M., Hida, K., Kohgo, T., Amemiya, A., Yoshida, K., and Fujinaga, K. (1996). Correlated expression of matrix metalloproteinases and ets family transcription factor E1A-F in invasive oral squamous-cell-carcinoma-derived cell lines. Am J Pathol *148*, 693-700.

Shou, C., Farnsworth, C. L., Neel, B. G., and Feig, L. A. (1992). Molecular cloning of cDNAs encoding a guanine-nucleotide-releasing factor for Ras p21. Nature *358*, 351-354.

Siegel, R., Naishadham, D., and Jemal, A. (2013). Cancer statistics, 2013. CA Cancer J Clin *63*, 11-30.

Singh, S. K., Clarke, I. D., Terasaki, M., Bonn, V. E., Hawkins, C., Squire, J., and Dirks, P. B. (2003). Identification of a cancer stem cell in human brain tumors. Cancer Res *63*, 5821-5828.

Soda, M., Choi, Y. L., Enomoto, M., Takada, S., Yamashita, Y., Ishikawa, S., Fujiwara, S., Watanabe, H., Kurashina, K., Hatanaka, H*., et al.* (2007). Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. Nature *448*, 561-566.

Soda, M., Takada, S., Takeuchi, K., Choi, Y. L., Enomoto, M., Ueno, T., Haruta, H., Hamada, T., Yamashita, Y., Ishikawa, Y*., et al.* (2008). A mouse model for EML4-ALK-positive lung cancer. Proc Natl Acad Sci U S A *105*, 19893-19897.

Soucek, L., Whitfield, J., Martins, C. P., Finch, A. J., Murphy, D. J., Sodir, N. M., Karnezis, A. N., Swigart, L. B., Nasi, S., and Evan, G. I. (2008). Modelling Myc inhibition as a cancer therapy. Nature *455*, 679-683.

Srinivas, S., Watanabe, T., Lin, C. S., William, C. M., Tanabe, Y., Jessell, T. M., and Costantini, F. (2001). Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. BMC Dev Biol *1*, 4.

Sullivan, J. P., Minna, J. D., and Shay, J. W. (2010). Evidence for self-renewing lung cancer stem cells and their implications in tumor initiation, progression, and targeted therapy. Cancer Metastasis Rev *29*, 61-72.

Sutherland, K. D., Proost, N., Brouns, I., Adriaensen, D., Song, J. Y., and Berns, A. (2011). Cell of origin of small cell lung cancer: inactivation of Trp53 and Rb1 in distinct cell types of adult mouse lung. Cancer Cell *19*, 754-764.

Suzuki, T., Sobue, T., Fujimoto, I., Doi, O., and Tateishi, R. (1990). Association of adenocarcinoma of the lung with cigarette smoking by grade of differentiation and subtype. Cancer Res *50*, 444-447.

Takahashi, T., Sonobe, M., Menju, T., Nakayama, E., Mino, N., Iwakiri, S., Nagai, S., Sato, K., Miyahara, R., Okubo, K.*, et al.* (2010). Mutations in Keap1 are a potential prognostic factor in resected non-small cell lung cancer. J Surg Oncol *101*, 500-506.

Tomlins, S. A., Mehra, R., Rhodes, D. R., Smith, L. R., Roulston, D., Helgeson, B. E., Cao, X., Wei, J. T., Rubin, M. A., Shah, R. B., and Chinnaiyan, A. M. (2006). TMPRSS2:ETV4 gene fusions define a third molecular subtype of prostate cancer. Cancer Res *66*, 3396-3400.

Tomlins, S. A., Rhodes, D. R., Perner, S., Dhanasekaran, S. M., Mehra, R., Sun, X. W., Varambally, S., Cao, X., Tchinda, J., Kuefer, R.*, et al.* (2005). Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science *310*, 644-648.

Trahey, M., and McCormick, F. (1987). A cytoplasmic protein stimulates normal N-ras p21 GTPase, but does not affect oncogenic mutants. Science *238*, 542-545.

Trahey, M., Wong, G., Halenbeck, R., Rubinfeld, B., Martin, G. A., Ladner, M., Long, C. M., Crosier, W. J., Watt, K., Koths, K., and et al. (1988). Molecular cloning of two types of GAP complementary DNA from human placenta. Science *242*, 1697-1700.

Travis, W. D. (2002). Pathology of lung cancer. Clin Chest Med *23*, 65-81, viii.

Travis, W. D., Travis, L. B., and Devesa, S. S. (1995). Lung cancer. Cancer *75*, 191-202.

Trimble, M. S., Xin, J. H., Guy, C. T., Muller, W. J., and Hassell, J. A. (1993). PEA3 is overexpressed in mouse metastatic mammary adenocarcinomas. Oncogene *8*, 3037-3042.

Vendramini-Costa, D. B., and Carvalho, J. E. (2012). Molecular link mechanisms between inflammation and cancer. Curr Pharm Des *18*, 3831-3852.

Visvader, J. E., and Lindeman, G. J. (2012). Cancer stem cells: current status and evolving complexities. Cell Stem Cell *10*, 717-728.

Vogel, U. S., Dixon, R. A., Schaber, M. D., Diehl, R. E., Marshall, M. S., Scolnick, E. M., Sigal, I. S., and Gibbs, J. B. (1988). Cloning of bovine GAP and its interaction with oncogenic ras p21. Nature *335*, 90-93.

Wallace, M. R., Marchuk, D. A., Andersen, L. B., Letcher, R., Odeh, H. M., Saulino, A. M., Fountain, J. W., Brereton, A., Nicholson, J., Mitchell, A. L., and et al. (1990). Type 1 neurofibromatosis gene: identification of a large transcript disrupted in three NF1 patients. Science *249*, 181-186.

Wang, Y., and You, M. (2001). Alternative splicing of the K-ras gene in mouse tissues and cell lines. Exp Lung Res *27*, 255-267.

Warburton, D., Schwarz, M., Tefft, D., Flores-Delgado, G., Anderson, K. D., and Cardoso, W. V. (2000). The molecular basis of lung morphogenesis. Mech Dev *92*, 55-81.

Wei, W., Mosteller, R. D., Sanyal, P., Gonzales, E., McKinney, D., Dasgupta, C., Li, P., Liu, B. X., and Broek, D. (1992). Identification of a mammalian gene structurally and functionally related to the CDC25 gene of Saccharomyces cerevisiae. Proc Natl Acad Sci U S A *89*, 7100-7104.

Wellbrock, C., Karasarides, M., and Marais, R. (2004). The RAF proteins take centre stage. Nat Rev Mol Cell Biol *5*, 875-885.

West, M., Kung, H. F., and Kamata, T. (1990). A novel membrane factor stimulates guanine nucleotide exchange reaction of ras proteins. FEBS Lett *259*, 245-248.

Worden, F. P., and Kalemkerian, G. P. (2000). Therapeutic advances in small cell lung cancer. Expert Opin Investig Drugs *9*, 565-579.

Xia, C. L., Zhu, P., Cai, Y. T., Zhu, G. B., Mei, Z. R., Huang, H., Luo, D. X., and Yan, P. K. (2012). HIV-infection resistance in PMBC-derived dendritic cells modified with recombinant virus. Arch Virol *157*, 413-421.

Xin, J. H., Cowie, A., Lachance, P., and Hassell, J. A. (1992). Molecular cloning and characterization of PEA3, a new member of the Ets oncogene family that is differentially expressed in mouse embryonic cells. Genes Dev *6*, 481-496.

Xu, G. F., O'Connell, P., Viskochil, D., Cawthon, R., Robertson, M., Culver, M., Dunn, D., Stevens, J., Gesteland, R., White, R., and et al. (1990). The neurofibromatosis type 1 gene encodes a protein related to GAP. Cell *62*, 599-608.

Xu, X., Rock, J. R., Lu, Y., Futtner, C., Schwab, B., Guinney, J., Hogan, B. L., and Onaitis, M. W. (2012). Evidence for type II cells as cells of origin of K-Ras-induced distal lung adenocarcinoma. Proc Natl Acad Sci U S A *109*, 4910-4915.

Yan, C., Ding, X., Wu, L., Yu, M., Qu, P., and Du, H. (2013). Stat3 Downstream Gene Product Chitinase 3-Like 1 Is a Potential Biomarker of Inflammation-induced Lung Cancer in Multiple Mouse Lung Tumor Models and Humans. PLoS One *8*, e61984.

Youssef, K. K., Van Keymeulen, A., Lapouge, G., Beck, B., Michaux, C., Achouri, Y., Sotiropoulou, P. A., and Blanpain, C. (2010). Identification of the cell lineage at the origin of basal cell carcinoma. Nat Cell Biol *12*, 299-305.

Zhang, Z., Wang, Y., Vikis, H. G., Johnson, L., Liu, G., Li, J., Anderson, M. W., Sills, R. C., Hong, H. L., Devereux, T. R.*, et al.* (2001). Wildtype Kras2 can inhibit lung carcinogenesis in mice. Nat Genet *29*, 25-33.

Zhou, G. B., Meng, Q. G., and Li, N. (2010). In vitro derivation of germ cells from embryonic stem cells in mammals. Mol Reprod Dev *77*, 586-594.